Title	A Study of Game Engagement through Gamification and Full-body Games
Author(s)	Silpasuwanchai, Chaklam
Citation	高知工科大学, 博士論文.
Date of issue	2015-03
URL	http://hdl.handle.net/10173/1277
Rights	
Text version	ETD

A Study of Game Engagement through Gamification and Full-Body Games

Chaklam Silpasuwanchai

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Engineering

Kochi University of Technology

2015

Supervisory Committee:

Xiangshi Ren

Hiroaki Shigemasu

Yukinobu Hoshino

Yoshiaki Takata

Masanori Hamamura

Abstract

Engagement is at the heart of successful interactive products. Failing to engage users can lead to less participation, less learning, less sales, and other undesirable outcomes. Consequently, a product is not realized to its full potential. As O'Brien and Toms [85] have concluded - *"Successful technologies are not just usable; they engage users."*

Nevertheless, engaging users remains an ongoing challenge for designers and researchers. Some basic research questions include "What is engagement?", "What are the associated factors?", "What is the effect of these factors on user engagement?" By first answering these fundamental questions, developers can begin to design better engagement.

This dissertation studies user engagement through the lens of digital games. A total of eight studies were conducted to identify and evaluate game engagement factors. The main results include (i) a development of an engagement framework composing of six-dimensional factors and (ii) an evaluation of these factors in design. Our work provides theoretical and practical foundations in academic research and design for, e.g., education, health, and entertainment.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express sincere appreciation to Kochi University of Technology who had given me the research opportunity.

I would like to thank my advisor, Xiangshi Ren, and express gratitude for his guidance and support.

I would like to thank my committee members, Yukinobu Hoshino, Hiroaki Shigemasu, Yoshiaki Takata and Masanori Hamamura for their valuable feedback throughout the project.

I would like to thank John Cahill for helping my writings.

I would like to thank the members of the Ren lab for their enormous support.

I would like to thank Aaron Marcus, Effie Law, Hirohide Yamada, Lennart Nacke, Mary Beth Oliver, Shinsuke Shimojo, and Xiaojuan Ma for their insightful comments.

I would like to thank my parents, Shing-Kwai Chong and Montip Silpasuwanchai, my brothers, Chak-Sam Chong and Hong-Choi Chong, and my fiancé, Napassaporn Likhitsajjakul for all their support.

DEDICATION

to my dear mother, Montip

Chapter 1

INTRODUCTION

Engagement is a key to the success of every interactive products. However, engaging users remains an ongoing challenge for designers and researchers. This is due to the fact that designs are often not adequately supported by a fundamental understanding of engagement. By first answering fundamental questions such as “What is engagement?”, “What are the associated factors?”, “What is the effect of these factors on user engagement?”, developers can begin to design better engagement.

Past works have associated engagement with several concepts, namely *aesthetics* [8] - how the products look and feel; *emotions* [67] - the ability of the products to evoke various emotions; *needs satisfaction* [30] - the ability to satisfy human basic needs; and *flow* [25] and *immersion* [58] - the ability to draw our attention into the tasks. But what is still missing is an integrated framework to holistically define what is engagement and how these different concepts relate to engagement.

To study engagement, we chose to explore digital games. We are motivated by the ability of games to deeply engage users and thus we considered digital games to be a logical starting point of investigation. Our main objective is to understand the notion of engagement, to identify the associated factors, and to evaluate the factors in design.

This disseration is structured as follows (see Figure 1.1). First, to identify and prioritize game engagement factors, we conducted two *large-scale investigation studies* - (i) a systematic review of existing engagement theories (Study 1) and (ii) a large-scale online survey study (Study 2). Second, to evaluate the effect of engagement factors, we conducted three *empirical* studies on gamificaton studying the application of engagement factors to learning (Studies 3-5), and three studies on full-body games studying player differences (Studies 6-8).

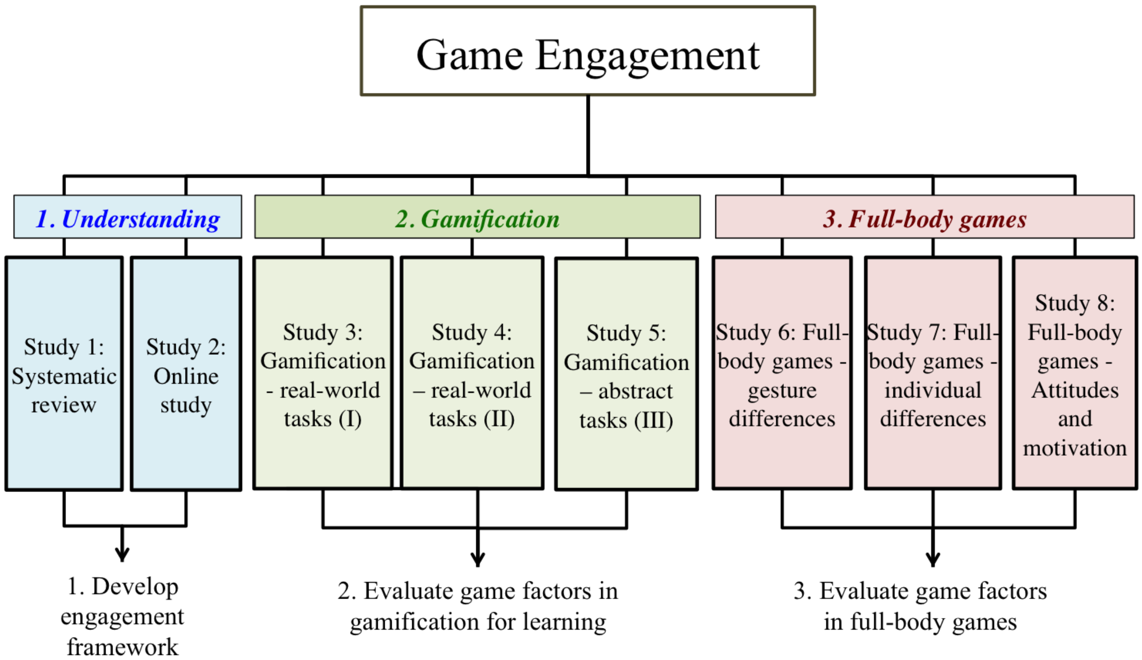


Figure 1.1: Dissertation outline.

Our main contributions include (i) a development of a integrated framework to define engagement from six different dimensions - *moderator*, *philosophical*, *psychological*, *behavioral*, *visceral* and *physical*, and (ii) the evaluation of these factors in design through empirical studies in gamification and full-body games. This work provides a theoretical and practical foundations in academic research and design for various application areas such as eduation, health and entertainment.

Chapter 2
RELATED WORK

We discuss (i) existing definitions of engagement, (ii) associated factors of engagement and (iii) the current research gap.

2.1 Existing definitions of engagement

Traditional definition of engagement refers to users' time-on-task or clicks-per-page in websites [47]. Traditional engagement also refers to the appearance and the attractiveness of products [47].

Modern definition of engagement shifts to a more holistic appreciation of engagement. Chapman [21] defined engagement as "something that draws us in, that attracts and holds our attention". Laurel [65] defined engagement as "a desirable, even essential, human responses to computer-mediated activities". Toms and O'Brien [85] defined engagement as "a quality of user experience and it may be embedded in a larger event or experience". Csikszentmihalyi [25] described a "flow" state where users are deeply engaged with the task-at-hand while forgeting everything around them.

More theoretically, Winograd and Flores [121] mentioned the notion of *engaged activity* as actions situated within our everyday activities. McCarthy and Wright [74] argued that meaningful engagement "depends on the event or action being felt, known, and varied in unique ways." Dewey [32] described the notion of *holistic engagement* where action, emotion, and thinking are all connected, with the "whole" person engaging with the activities.

This review shows the variety of definitions across domains, in which there is no agreement upon. One objective of this dissertation is develop a integrated definition of engagement in a form of framework.

2.2 Associated factors of engagement

Past works have associated engagement to many factors. Engagement has been associated with product's aesthetics [8] as it appeals to our sensory organs. These aesthetical factors include product's apperance, sound, tactile properties, and olfactory properties.

Engagement has been associated with emotions and affect [67]. Lazzaro [67] identified over thirty emotions, e.g., joy.

Engagement has been associated with product's usability, reliability, utility and performance [48].

Engagement has been associated with needs satisfaction [30]. Self-Determination Theory [30] identified autonomy, competence, and relatedness as three basic drives of motivation.

Engagement has also been associated with cognitive dissociation, namely flow [25] and immersion [58]. In a flow state, one is completely absorbed with the tasks-at-hand. Csikszentmihalyi [25] further describes eight flow factors: clear goals, balance between skills and challenges, merging of action and awareness, concentration, autonomy, loss of consciousness, time distortion, and autotelic experience.

Past works have identified multiple factors of engagement but they are disconnected conceptually. One objective of this dissertation is to synthesize these factors and to understand the relationship of these factors with engagement.

2.3 Summary of research gaps

Our review revealed two important research gaps regarding the study of engagement. First, although there has been some works on engagement, they are often fragmented, scattered and often repetitive. There is a need to synthesize these bits and pieces of knowledge to develop a more holistic, integrated understanding of engagement. Second, most evaluations of engagement factors are not adequately grounded in theory. Without a strong theoretical foundation, it is difficult for designers and researchers to reproduce engagement in a reliable way. This dissertation seeks to address these two research gaps.

Chapter 3

STUDY 1: SYSTEMATIC REVIEW OF GAME ENGAGEMENT THEORIES

This study aims to develop a integrated framework for engagement which is grounded in theory. We used *systematic review* methdology to review existing game engagement theories and to identify common themes. There are many theoretical studies on game engagement scattered across different disciplines, but they have not been organized into a focused study of game engagement. By synthesizing and organizing this information, a comprehensive account of game engagement can be developed. We developed a integrated theoretical framework to describe the notion of game engagement. We also discussed the relationships between engagement and related concepts, and related game engagement features.

3.1 Methodology

We seek to understand human engagement in digital games, particularly via theoretical literatures that provide explanatory power on how games engage users. To achieve that, we conducted a survey where *systematic review* methodology was used. The methodology consists of four steps: (1) Identifying relevant databases and collections; (2) identifying relevant search terms/keywords, (3) specifying selection criteria and (4) performing coding analysis.

3.1.1 Databases and Data Collection

We searched databases that are commonly used for publications in digital games, including those identified as relevant to information technology, psychology, and social sciences: ACM Digital Library, DiGRA, Web of Science, Scopus, IEEExplore, Science Direct and EBSCO.

We restricted our search from 2000 to 2014, given the surge of interest in game studies since 2000 [15] as seen in the establishment of game conferences such as ACE, Future Play, DiGRA, or SIGGRAPH Sandbox in the last decade (some of them recently merged into the new CHI PLAY conference series). To prevent possible loss of important theories before 2000 or other sources of information, we used the snowballing method [57], where commonly cited papers shared across relevant documents were also collected, thus we were able to retain important references, such as Csikszentmihaly's Flow theory [25] and Caillois's [19] and Huizinga's Play theory [51]. While we are aware that any literature search method is bound to omit some relevant work, we believe that our method allowed us to include the most relevant references on our subject.

3.1.2 Search Terms

We derived search terms from previous reviews [15, 76]. Since our goal was to scan the references explicitly for the impact of games in engaging or motivating players, we added new search terms for this. Our resulting search terms include: "Games," "Engagement," "Enjoyment," "Fun," "Motivation," "Attention," "Emotion" and "Affect."

3.1.3 Selection Criteria

Our search terms resulted in many retrieved papers from the databases. To keep things manageable, we narrowed the scope for inclusion of papers using the following three screening criteria: The documents had to (1) provide a *theoretical contribution* regarding game engagement; this includes theoretical papers or papers that were linked to existing theories. We discuss some general human engagement theories (e.g., Self-Determination Theory) to provide additional explanatory power. Second, due to the primary focus of past works on theorizing game engagement in digital entertainment games, we focused on documents that (2) discuss game engagement in *digital entertainment games*; thus serious games or gamification-related studies were omitted. Although we regard serious games and gamification as important emerging fields, they were out of the scope of this review. Finally, the documents should be (3) original, peer-reviewed papers written in the English language. Using these three limiting conditions a total of 17 theories and 74 related documents were included for the current review (see Table 3.1).

3.1.4 Coding Analysis

Each document was coded according to: theory/concept, author name(s); published year; field of study; presented factors associated to game engagement; related evidence (approach, measures) (see Table 1 for overview). Our coding analysis show that most theories originated around 1950s to 1990s and it was until recent decade that these theories have been investigated in digital games. The original source of theories includes psychology, media psychology, social psychology, games, PX (player experience), and cultures. Presented factors associated to game engagement include needs, emotions, flow, immersion, presence, audiovisuals, embodiment, realism, attitude, goals, and relationships. The most common methodologies used to investigate these theories in games are empirical lab study, field study and large-scale survey study. Participants averagely ranged from age 20 to 30. The most common method for measuring engagement is questionnaires; other measures include physiological measures and performance-related metrics.

3.2 Classification Framework

Our intention is to synthesize multiple theories from different sources which are fragmented, scattered and often repetitive. Without a high-level framework, it can be difficult to understand the position of each theory. To achieve this, we conducted an grounded-theory analysis [44] on our collection of theories and studies. Each theory and study was coded and categorized with keywords. The analysis revealed four high-level themes associated with game engagement (i) needs satisfaction, (ii) emotion/affect, (iii) cognition, and (iv) relationships. These emerging themes were confirmed by an independent rater with high inter-rater reliability (Kappa=0.926, $p<0.01$). We used this framework to classify existing game engagement theories. The classification was based on descriptions in Table 3.2:

Theory	**Authors**	**Field**	**Presented Factors**	**Related Evidence in Games**	**Approach, Participants**	**Measures**
Theme 1: Needs Satisfaction						
SD theory	(Deci & Ryan 2000)	Psychology	Needs	Ryan et al. [102] found that autonomy, competence and relatedness predict enjoyment and long-term engagement	3 lab studies and 1 survey study; Total n>100; mean age =22.1	Questionnaires
U&G Theory	Katz & Blumler, 1974	Media Psychology	Needs	Sherry et al. [106] analyzed the uses and gratifications of games	Survey study; n>100; mean age=16.93	Questionnaires
Theme 2: Emotion						
Emotion Theory	Various	PX (Player Experience)	Emotions	Lazzaro [67] identified emotions in games	Survey study; n=30; Mean age = not specified	Questionnaires, Interviews, Observation, Facial analysis
Play Theory	(Caillois, 1961; Huizinga, 1950)	Culture	Play	Caillois [19] observed and classified play in traditional games such as poker and board games	Conceptual	-
Mood Management Theory	(Zillman & Bryant, 1985)	Media Psychology	Emotions; Selective exposure	Reinecke et al. [97] found that games can repair mood	Lab study; n>100; mean age=19.96	Questionnaires
Affective Disposition Theory	(Raney, 2003)	Media Psychology	Basal morality	Klimmt et al. [62] studied disposition (suspense) in games	Lab study; n=63; mean age=20.6	Questionnaires
Theme 3: Cognition						
Flow Theory	(Csikszentmihalyi, 1990)	Positive Psychology	Flow	Sweester and Wyeth [111] evaluated GameFlow Model	Expert evaluation, 32 reviews	Heuristic evaluation
Immersion	Various	PX	Immersion	Jennett et al. [58] studied immersion in games and developed immersion questionnaire	Three lab studies; total n>100; mean age=23.66	Questionnaires, Eye movements, Task completion time
Presence	Various	Media Psychology; PX	Presence	Slater et al. [109] studied depth of presence in virtual environment	Lab study; n=24; mean age=not specified	Questionnaires
Attitude	(Ajzen, 1985; Davis, Bagozzi, & Warshaw, 1989)	Psychology	Attitude	Wu and Liu [124] studied the effect of attitude in online games	Survey study; n=253; mean age=23.2	Questionnaires
Embodiment Theory	Various	Media Psychology; PX; Whole-body interaction	Embodiment	Bianchi-Berthouze [11] explored how body movement affects player engagement	Three lab studies; total n=38; mean age=22.5	Questionnaires; performance-related metrics
Realism and Fidelity	Various	Media Psychology; Simulation; PX	Fidelity	Ivory & Kalyanaraman [56] found the increased realism can enhance sense of presence in games	Lab study; total n=120; mean age=20.57	Questionnaires; physiological measures (skin conductance)
Audiovisual	Various	PX	Audiovisuals	Nacke et al. [81] found the absence of sound affects game engagement	Lab study; n=36; mean age =24	Physiological measures (EMG, EDA); Questionnaire
Theme 4: Relationship						
FIRO Theory	(Schutz, 1958)	Social Psychology	Relationships	Lucas & Sherry [71] studied FIRO theory in games	Survey study; n>100; mean age=19.71	Questionnaires
Social Facilitation Theory	(Zajonc, 1965)	Social Psychology	Relationships	Kort et al. [28] studied social presence and developed SPGQ	Survey study; n>100; Mean age=19.8	Questionnaires
Social Comparison Theory	(Festinger, 1954)	Social Psychology	Relationships	Ryan et al. [102] showed the impact of game experience on self-esteem	3 lab studies and 1 survey study; Total n>100; mean age = 22.1	Questionnaires

Table 3.1: Overview of theories and concepts in this review.

Themes	**Description**
Needs Satisfaction	Players innate nature and fundamental psychological needs are in focus here. Common questions include: How do games satisfy players' psychological needs? What are players psychological needs?
Emotion/Affect	Players emotional responses and feelings are affective factors. Common questions are: What is the impact of games on emotions? How does positive and negative valence impact players' engagement?
Cognition	Players thought process, perception, embodiment and visual attention affect game engagement. Common questions include: What is the effect of games on players cognition (e.g., loss of awareness)? How does one reach those states? How do games visually and audibly affect players?
Relationships	The impact of social influences on players engagement is in focus here. Common questions include: What governs a positive and negative social experience? How do social experiences impact players engagement as a whole?

Table 3.2: The classification framework of engagement theories.

3.3 Theme 1: Needs Satisfaction

This theme facilitates understanding of game engagement by discussing the characteristics of human nature. The following theories provide theoretical explanations regarding this theme.

3.3.1 Self-Determination Theory

In Self-Determination theory (SDT) [30], Deci and Ryan classified motivation into two categories: *intrinsic* motivation, driven by inner needs, and *extrinsic* motivation, driven by external factors, such as rewards or threats. They described intrinsic motivation as drives, which all human-beings will strive to meet: their inner drives for *autonomy*, *competence*, and *relatedness*. The feeling of independence, being in control of things, feeling the "*origins*" of their own actions, and making their own choices are primary inner drives for *autonomy*.

Competence or *mastery* is the drive to fulfill one's need of feeling in control through skill mastery. Lastly, *relatedness* is the need to connect, to interact, to be accepted, and to be understood. While human beings prefer to be independent and competent, they are also motivated by acknowledgement from others, regarding their independence and competence.

Ryan et al. [92, 102] argued that SDT theory can be used to explain the underlying motivational structure of video games, because the theory focuses on human basic needs, which are found across games and player types. Through a series of studies, Ryan et al. [102] found that the SDT (cf., Player Experience of Need Satisfaction [PENS] model) can be used to predict enjoyment and long-term engagement. Similar results were presented in related literature [115].

3.3.2 Uses and Gratification Theory

Uses and gratifications (U&G) theory [59] describes that people engage with media (e.g., games) to satisfy their specific needs (e.g., enhancing knowledge, escape, relaxation). In addition, U&G theory argues that every individual has different needs based on their past experiences, interests and motives. For example, some people may engage with a game for relaxation, while others may play games to fulfill their need of feeling competent. Finally, U&G theory assumes that players are active audiences, who have control over what games they would play, suggesting that game engagement is voluntary and selective. Using U&G as a theoretical foundation, researchers have identified several uses of video games (e.g., to relax/escape/kill time/avoid doing other things [89, 106], to compete [106, 125], to achieve [72, 125], to socialize [106, 125], to be aroused [39, 67], to explore/discover/learn [72, 125], and to fantasize [72, 106]).

3.3.3 Discussion

This theme has suggested that games are engaging because they are able to satisfy players' psychological needs, such as the need to feel autonomous, competent, and socially-connected.

There are evidence [102] that games satisfy the need of autonomy, competence , and relatedness. There are also evidence that games satisfy other needs such as escape and relaxation [106]. Researchers found that needs satisfaction predicts long-term engagement [102, 115]. Among all the needs, challenge (competence) is consistently rated as the key factors for engagement in SD Theory [102] and in U&G Theory [106].

Self-Determination Theory is often criticized for identifying narrow range of needs and U&G Theory identified a broader range of needs, e.g., escape, relaxation. It seems that other needs will be identified in the future. U&G Theory has also stated that needs vary in extent across people based on their past experiences, interests and motives but there has been lack of integrated understanding of how exactly these needs vary across person.

U&G Theory has informed that needs satisfaction is a selective and voluntary process. What intrigues us is why games may be chosen over other media or activities as other activities may also equally provide the same type of needs satisfaction. This leads to the recognition that features of the games themselves also contribute to game engagement. OBrien and Toms [85] associated these features with "engagement attributes" including interactivity, perceived control, and novelty.

It has been stated that needs satisfaction may in part explain how game engage users. Oliver and Raney [86] used the word "eduaimonic" to link these needs pertaining to well-being, purpose and meaningfulness. The evidence that players are purpose-seeking beings can be reflected from players' reported motives [106, 125] including the motives to win, to make progress, to interact with others, to explore, discover, and learn.

3.4 Theme 2: Emotion

This theme provides a means of understanding engagement through the lens of emotions.

3.4.1 Emotion Theory

Researchers [39, 67, 96] have suggested that games are engaging because of the ability of games to evoke many different emotions. Ravaja et al. [96] defined emotion as *"biologically-*

based action dispositions that have an important role in the determination of behavior." Emotion contains three dimensions: (1) *subjective experience* (e.g., feeling happy), (2) *expressive behavior* (e.g., smiling), and (3) *physiological activation* (e.g., sympathetic arousal). Emotion (especially for physiological game assessment studies) is often described using a two-dimensional circumplex model [101] with two factors — valence (negative or positive emotions) and arousal (intensity of the emotions).

Ravaja et al. [96] stated that games are successful, because they are able to elicit a wide range of strong emotional responses, from fun and satisfaction to guilt and sadness. Lazzaro [67] found over thirty different emotions that make games fun and enjoyable. Lazzaro synthesized these emotions info four different types of fun: Hard Fun (Frustration and Fiero), Easy Fun (Wonder, Curiosity), Serious Fun (Excitement, Relief) and People Fun (Recognition and Connection). Hunicke et al. [52] identified eight types of fun: sensation, fantasy, narrative, challenge, fellowship, discovery, expression and submission. More broadly, sensitivity theory [98] defines fun (or joy) as the satisfaction of 16 basic human desires (e.g., curiosity, power), with each of these desires linked to a particular emotion.

Enjoyment is a term that has been often associated with positive affect and identified as an affective outcome of a good gaming experience [117]. Enjoyment is consistently regarded as key explanation for game engagement. Mekler et al. [76] described enjoyment as the valence (affective aspect) of the player experience (fun, interest, pleasures). Blythe and Hassenzahl [13] described enjoyment in the dimension of distractions.

3.4.2 Play Theory

In Huzinga's play theory [51], he argued from a cultural perspective that "play" is essential to all human beings stating that *"play is older than culture [...] all culture is an element of play."* He further added that the most significant aspect of play is fun. Caillois [19] further described four fundamental types of play: Agôn (competition), Alea (chance and uncertainties), Mimicry (role-playing), and Ilinx (changing state of mind and perception). He further described play along a dimension of interactive freedom. Ludus being a rule-based form of play and paida being free-form improvisational play. Voluntary play or freedom of play (without third-party purpose) is a core aspect of games. One possible explanation of why humans enjoy the voluntary nature of play is that gamers are able to interact and express themselves more freely and emotionally, while not feeling controlled or monitored [19].

3.4.3 Mood Management Theory

Based on the assumption that humans are pleasure-seeking beings, mood-management theory [128] states that to maximize pleasures, humans instinctively tend to expose themselves to favorable environmental stimuli such that positive valence (pleasures) is maximized, whereas negative valence (pain) is minimized. This theory is linked to selective exposure theory [128], which states that humans possess tendencies to expose themselves to information that reinforces their previous beliefs or views while avoiding contradictory information. Nevertheless, this theory does not address why some players engage with negatively valenced activities, such as scary interactions in horror games, where fear and suspense are the primary emotions. One explanation was proposed by Klimmt [61], which argued that players may engage with games that elicit negative feelings, because they anticipate a resolution that will not only alleviate the negative feelings, but will result in feelings of euphoria and a great sense of achievement.

3.4.4 Affective Disposition Theory

Affective disposition theory [95] states that players make dispositional judgment of and emotional reactions to characters in the media/virtual world, which in turn affects their pleasures and enjoyment. For example, players have a tendency to share the sympathy and hope of the main character, while to hope for a negative outcome for the villain. The theory suggests that love (for the hero) and hate (for the villian) are two strong emotions that makes story enjoyable and engaging.

Disposition theory may provide an explanation for why "Role-playing games" (RPGs) could be particularly enjoyable, as the player him/herself has the opportunity to facilitate negative outcomes for disliked characters, and to directly enjoy the victories of the "liked" protagonists. Indeed, researchers found that "Role-playing games" (RPG) as one of the most immersive type of game [112, 125]. Role-playing games are those, in which players assume the *role* of a character in the game, go through the story of the game as if they were there, and where, in many cases, their actions impact the ending of the story. The interaction between "characters" and the "story" is critical for RPGs because it allows games to evoke a wide range of strong emotions, which make games engaging [39, 112, 125]. In addition, the combination of a story with frustrations, dilemmas, decision making and multiple paths enables players to experience a deep level of emotional engagement and purpose [39, 55].

3.4.5 Discussion

This theme offers the theoretical perspective that games are engaging because of their capability to evoke a wide range of strong emotions both positive and negative. There are evidences that games elicit strong, wide range of emotions [96]. Play Theory, Mood Management Theory, and Affective Disposition Theory have provided theoretical explanations why humans are attracted to pleasures and emotional arousals from different perspectives (e.g., culture, mood, disposition). They also provided some mappings to game features that contribute to strong emotions including uncertainties, difficult challenge, role-playing/story, competition, and dilemmas.

Enjoyment is a term associated with positive affect and is consistently rated as key explanation for game engagement [76]. However, engagement is also associated with negative affect [58] suggesting that engagement may not occur due to enjoyment only but negative arousals such as suspense, guilt, frustrations [117]. Further research should investigate more in detail how negative emotions impact game engagement and how games can be designed to elicit these emotions. It would also be beneficial to understand how emotion should be best designed whether designers should design to elicit wide range of similar emotions or to elicit wide range of different emotions. Based on sensitivity theory, further research should also investigate how negative emotions relate with players needs satisfaction, desires, values and individual differences.

3.5 Theme 3: Cognition

This theme describes game engagement from the perspective of cognition, e.g., awareness, consciousness, attention.

3.5.1 Flow theory

Researchers [24, 107, 120] defined flow as the cognitive aspect of experience (involvement) with the task. In Csikszentmihalyi's flow theory [25], flow was defined as the mental state of being "*completely* immersed," losing complete awareness even of bodily needs, with all his/her attention completely dedicated to a particular task at hand. He added that flow occurs when there is an optimal alignment between a user's skill level and the challenges posed by the task. There are seven additional elements needed to support that optimal alignment: clear goals, merging of action and awareness, concentration, autonomy, loss of consciousness, time distortion, and autotelic experience.

Past works (e.g., [91]) have supported flow theory by showing that the most satisfying and engaging moment for players is when there is an optimal alignment between the player's level of skill and the challenges provided by the game (e.g., barely victorious), while, when the challenge is too easy or too difficult (e.g., totally victorious) for players, the game becomes less engaging. Flow theory is widely-accepted by researchers because of its universal nature, and thus it has been used extensively for explaining the phenomenon of game engagement (e.g., GameFlow model [111]). Nevertheless, Nakamura and Csikszentmihalyi [82] found that enjoyment may occur independently of flow (i.e., flow describes a extreme gaming experience, which may not cover more casual experiences of enjoyment and light-weight absorption).

3.5.2 Immersion

Jennett et al. [58] considered immersion as a result of good gaming experience. Immersion [18, 22, 58, 103] comprises three main features: (1) temporal dissociation, (2) spatial dissociation and (3) merging of task and self. Jennett et al. [58] argued that immersion is different from flow in the sense of extremity (i.e., since immersion is a less extreme version of flow). Thus, immersion can be used more effectively to describe a variety of player experiences (e.g., in casual gaming). Brown and Cairns [18] defined immersion as the degree of involvement within gameplay, ranging from low (engagement) to moderate (engrossment) to high immersion (total immersion). Douglas and Hargadon [33] viewed immersion as one of the primary sources of pleasures.

Ermi and Mäyrä [35] proposed three types of immersion: sensory, challenge and imaginative immersion. Calleja et al. [20] proposed a player involvement model composed of six types of involvement that facilities immersion: kinesthetic, spatial, shared, narrative, affective and ludic involvement. McMahan [75] proposed three important conditions for immersion: players' expectation should match the game conventions, meaningful play, and a consistent game world. Jennett et al. [58] found that not only positive affect, but negative affect such as anxiety may also promote immersion.

3.5.3 Presence

Presence is closely related to immersion. Most commonly, researchers [109, 120, 122] define presence as the sense of being there in the virtual environment without actually perceiving the existence of the medium. Researchers [109, 122] have found that the naturalness of the interactions and realism affect presence. Banos et al. [5] found an association between emotion and presence, i.e., affective content (story) increases presence in a virtual simulation role-playing game. Jennett et al. [58] argued that presence is only small part of gaming experience, e.g., one may experience immersion without presence in a puzzle game. On another hand, one may experience presence without immersion such as performing a boring task in a virtual simulation world. It can also be arguably said that presence is synonymous with spatial immersion [120].

In terms of the impact of presence on enjoyment, Lombard and Ditton [70] suggested that a high sense of presence leads to greater enjoyment. On the other hand, Weibel and Wissmath [120] found that presence does not directly impact enjoyment, but rather facilitates flow or immersion which in turn results in enjoyment (i.e., presence precedes flow). Weibel and Wissmath [120] also stated that the impact of presence is dependent on the types of games (i.e., presence is more important in vivid, realistic games like first-person shooters/role-playing games than in abstract puzzle/memory games, which require less realism).

3.5.4 Attitude

Researchers have found that attitudes may affect one's engagement [80]. Attitude [80] is defined as the psychological tendency to favor or disfavor certain entity. Two important theories of attitudes are the Technology Acceptance Model (TAM) [27] and the Theory of Planned Behavior (TPB) [2]. TAM considers perceived usefulness and ease of use as two predecessors determining attitudes, which affect their engagement. Lee and Tsai [68] replaced perceived usefulness with perceived enjoyment when examining the effect of attitudes on online games. TPB states that attitudes together with a perceived social norm and behavioral control shape one's intentions, which in turn shape the final behaviors. By viewing attitudes as individual characteristics, we may better understand how behavioral outcomes can also differ based on individual differences and external factors (social norm).

3.5.5 Embodiment Theory

Researchers [7, 10, 42] described game experiences as an embodied phenomenon. Embodied cognition [42] refers to where mind and body are connected and how they influence one another, specifically arguing that bodily experiences can influence cognitions, unlike previous assumptions of the mind and body as separate entities. The concept of embodiment has often been used by researchers to describe the experience in full-body games [11] which have shown that body movement influences ones emotions and engagement. Embodiment also has been used to describe some role-playing games [7, 10, 42], where the player's mind is influenced by the avatar's bodily experiences (i.e., the concept of embodiment can explain how players can become one with the avatar and feel deeply immersed). Embodiment illuminates our understanding that the mind, the body and the environment (input device, outdoor, indoor) are all connected, which influences the player's engagement. This entirely suggested that cognition (embodied cognition) is not solely composed of the mind, but also influenced by bodily actions.

3.5.6 Realism and Fidelity

Realism is the extent to which a game resembles the real world. Realism is affected by the quality of aesthetics (visual and audible) in games, as well as the surrounding environment of players during gameplay. A similar term is fidelity, which Hays and Singer [49] defined as the "degree of correspondence between simulation and real circumstances", i.e., fidelity may cover broader scope of realism to include physics and natural laws. Often, the more realistic the game, the more easily players feel a higher sense of presence, and more easily become immersed in the game, especially in vivid and realistic games (e.g., first-person shooting or simulation game) [49, 75, 109, 122] . Several studies about realism have been conducted, e.g., artificial gun vs. mouse [60]; large screen vs. PC monitor [5]; stereoscopic 3D games vs. 2D games [104]. These studies indicate that realism increased level of presence, however Weibel and Wissmath [120] implied that realism plays a more important role in vivid, realistic games (first-person shooting/role-playing game) than in other puzzle/abstract games which requires less realism.

3.5.7 Audiovisual

Music and sound engage users by evoking and enhancing the intensity of emotions [99]. Parker and Heerema [87] described that sound creates a feeling of presence, reminding gamers that the game is still going on. Fast music may represent a lot of activity, and vice-versa for slow music. They also suggest that sound affects emotions faster than visual display. Nacke et al. [81] found significant correlations between audio and game engagement constructs.

In terms of visuals, LaViola and Litwiller [66] found that players enjoyed playing using a 3D stereo display compared to a 2D display. Ermi and Mäyrä [35] found that audiovisual capability and visual-motor links are fundamental in enabling a higher quality of gaming experience such as immersion. Takatalo et al. [114] reported that screen size has no significant impact on engagement, although Banos et al. [5] found otherwise. It appears that the importance of visual fidelity depends on the type of games (more important in role-playing/first person games). In terms of graphical aesthetics, Andersen et al. [3] found that gameplay variations affected play time three times as much as a variation in aesthetics. This finding suggests the supporting role of aesthetics on the overall gameplay.

3.5.8 Discussion

Game engagement has been associated with spatial and temporal awareness, described by the concept of flow, immersion and presence. While flow describes optimal experience, Jennett et al. [58] argued that immersion is a more useful concept than flow as it can be used to explain more casual gaming experience. Presence was often referred as teleportation to a virtual environment and may occur independently of immersion. Flow and immersion can be both seen as the motives of playing games [125], as well as a cognitive outcome of a good gaming experience [91].

Nakamura and Csikszentmihalyi [82] found that enjoyment may occur independently of flow. Mekler et al. [76] stated that enjoyment is different from flow, i.e., enjoyment is a characteristic of flow, but enjoyment may occur independently of flow. One may view enjoyment as the affective aspect of the gaming experience [117], while absorption (flow, immersion) as the cognitive aspect (involvement) of the experience [91].

Embodiment Theory has provided an interesting angle of engagement. There are evidence that bodily interactions with the environment affect engagement and performance [11]. This lead to other game features that could contribute to game engagement including the design of avatar, the game controllers and the physical environment. Embodiment might also provide explanation regarding the difference in engagement between physical board/card game and its virtual counterpart.

One aspect of engagement concerns audiovisual level of experience. There are evidence that audio and visual cues facilitate game engagement [81]. When they are not carefully designed, they may cause disruption in absorption or lower enjoyment. Realism is seen by researchers as important in realistic games (e.g., Role-playing) but not necessary in other types of games (e.g., puzzle game).

3.6 Theme 4: Relationships

The social perspective is concerned with the study of social factors as well as social diversity (female vs. male). In short, because human-beings seek acknowledgement from others for their competence and uniqueness, social experiences such as competition, collaboration and connection can be utilized to further facilitate and enhance game engagement.

3.6.1 FIRO Theory

Fundamental Interpersonal Relationship Orientation (FIRO) Theory [105] argues that all humans are governed by three social needs: inclusion, affection and control. Inclusion refers to the need to belong to a social group and the need to interact with others. Affection refers the need to feel the sense of love and warmth in relationships. Controls refers the need in which ones wants to have influence/control over others' decisions/actions. FIRO theory also asserts that these orientations and priorities vary across people. Lucas and Sherry [71] argued that these three factors can be exploited to enhance game engagement. For example, structuring gameplay around teamwork and collaboration satisfies players' need of inclusion, affection and control.

3.6.2 Social Facilitation Theory

Social facilitation theory [127] states that people have a tendency to perform differently in the presence of other people. Specifically, with presence of other people, one would perform better in well-familiar tasks, while perform worse in less-familiar tasks. Researchers mostly agreed that the changed performance is a result of awareness of possible evaluations from others, which can be readily observed in competition or collaboration scenarios in games. Nevertheless, the degree to which a person is influenced by social presence varies. Kort et al. [28] developed the SPGQ (Social Presence in Gaming Questionnaire) with social presence of others (e.g., playing with friends) as the primary source of motivation in gameplay. Other relevant theories include Social Proof Theory [23] which predicts that players are likely to engage in behaviors that others are also engaged in, while Social Identity Theory [113] states that humans share a sense of who they are based on their social groups (e.g., countries, gender, affiliations) as a process of self-image enhancement. Related to identity, Beenen et al. [9] suggested that individuals are most socially motivated when their uniqueness and contribution is being acknowledged in a team environment. Entirely, these theories suggest that humans are social in nature and that they seek the approval and avoid the disapproval. As a result, social factors such as competition, collaboration, social identity and status can be argued to play important roles in enhancing gameplay.

3.6.3 Social Comparison Theory

Social Comparison Theory [37] states that social experiences are driven by the need to better understand the self (accurate self-evaluations) as well as the need to improve one's self-esteem. This relationship between self evaluations, comparison with others, and self-esteem implies that needs pertaining to competence and relatedness in self-determination may be associated. For example, in an online game environment, players, driven by the need to improve their self esteem, may seek self-enhancement and verification from others about their skills level. If this observation is correct, it also implies that social experience may also partly driven by the need of competence. Thus social mechanisms, such as pushing high scores to the leader board, sharing trophies on public Web space, or even showing off their skills in public or with their friends, may further promote sense of competence. Nevertheless, some researchers indicate that social interaction may also possibly disrupt immersion and flow [111] and thus should be carefully designed to avoid degradation of the sense of competence or feelings of alienation. In addition, social interaction [111] can also break the link between real-world and fantasy world because real-people in social interactions can provide a link back to the real-world.

3.6.4 Discussion

Relationships can be viewed as one of needs satisfaction in SD Theory (Relatedness). Relationships is also closely associated with the feeling of competence [37], self-esteem [102] and feeling in control [105]. This association was reflected in experiments where social presence of others affect ones engagement and performance [127]. There is also evidence that humans seek acknowledgement from others for their uniqueness [9] and competence [102]. Further research should include investigating how different types of social presence (e.g., physical friends, online friends) affect game engagement.

Yee [125] identified socializing along with achievement and immersion as important motives for engaging games. In online games, socializing is identified as the key reason for playing [102] suggesting that socializing may be more important in some games. Some common game features contributing to social engagement include teamwork, communication channel, competition, and leaderboard.

Although relationships can improve player engagement, it was implied to carefully design social features such that the feeling of alienation or degradation of self-esteem should be avoided or minimized [37]. It should also be careful not to design social features that disrupt immersion or flow as real-people in social interactions may provide a link back to the real-world [111].

3.7 General Discussion

We discuss two issues: (1) what is game engagement and (2) relationships between game engagement and other concepts.

3.7.1 What is Game Engagement?

From the review, we can better understand about game engagement. Game engagement has been often associated with **needs satisfaction** in SD Theory and in U&G Theory. Needs satisfaction is considered by many researchers to be the key explanation of game engagement, where various psychological needs have been identified and mapped. SD Theory describes explicit, high level needs including autonomy, competence and relatedness, while U&G covers broader range of needs including implicit needs such as relaxation and pleasures. Game engagement has also been identified as a selective and voluntary process in U&G Theory, suggesting that game engagement varies across different persons.

Game engagement has also been associated with **emotions**. Particularly, some has treated game engagement in the synonymous fashion as enjoyment (positive affect) while some treated enjoyment as a key explanation of game engagement. Some also treated enjoyment as an affective component of game engagement. In any case, researchers found that engagement can occur in a negative-valenced gameplay (e.g., horror gameplay) suggesting that game engagement is associated with both positive and negative affect.

Game engagement has also been discussed in the dimension of **awareness**, absorption and distractions, namely the concept of flow, immersion and presence. It has been stated that when a player is engaged, they can achieve flow, immersion or presence a state where their awareness is dissociated spatially and temporally. Flow Theory stated eight components to achieve the state of flow clear goals, merging of action and self, concentration, autonomy, loss of consciousness, time distortion, autotelic experience, and alignment between challenge and skills.

In the cognitive level, game engagement has been described as an **embodied phe-**

nomenon. Our mind or cognition is affected by how we act on the environment and thus suggesting that game engagement is affected by bodily interactions with the environment, e.g., game controller, avatar, physical environment. Designing game engagement thus will also need to consider the medium and environment of gameplay.

Game engagement has also been associated with our **senses** how something looks and feels. These visceral features (audiovisuals, realism) impact our initial engagement and may disrupt our level of immersion and enjoyment when not correctly designed.

In the social level, game engagement has been associated with **relationships** as seen in many online games. Because players are motivated by the need to connect and to be approved by friends, social interaction impacts whether and how long a player will engage in a game. Relationships have also been closely related with the sense of feeling in control, self-esteem and competence.

While the aforementioned factors can be used to explain game engagement, they may occur outside of gameplay, e.g., watching movies. This leads to the recognition that **game features** themselves also contribute to game engagement. These features are distinct from other types of activities/media which makes game engaging.

Entirely, researchers have agreed that game engagement is a complex, multi-faceted phenomenon composing of multiple associated factors needs, emotions, awareness, relationships, game feature, and individual differences. We propose that associated factors of game engagement may be viewed as six-dimensional (we called Six-Dimensional Engagement Framework) as shown in Figure 3.1.

The *moderator* dimension describes how each person is fundamentally different based on their beliefs, cultures, attitudes and past experiences. These differences determine different values, different priority of needs, different gaming motives and different gaming preferences.

Game engagement can be viewed from the *philosophical* dimension, understanding human values and paradigms. Our review suggests that human is pleasure-seeking beings (seeking to have fun), purposeful-seeking beings (seeking to learn and master) and social beings (seeking to connect with other people).

From the *psychological* dimension, game engagement can be understood from what humans needs. Our review shows that human is satisfied by the basic needs of autonomy, competence and relatedness. Human is also satisfied by other more implicit needs such as emotional arousals and relaxation.

From the *behavioral* dimension, game engagement can be experienced cognitively as absorption and affectively as enjoyment or suspense.

The *visceral* dimension describes game engagement from the sensory level including appearance, sound and realism.

The *physical* perspective describes games features, attributes, interaction and the physical environment which contributes to game engagement.

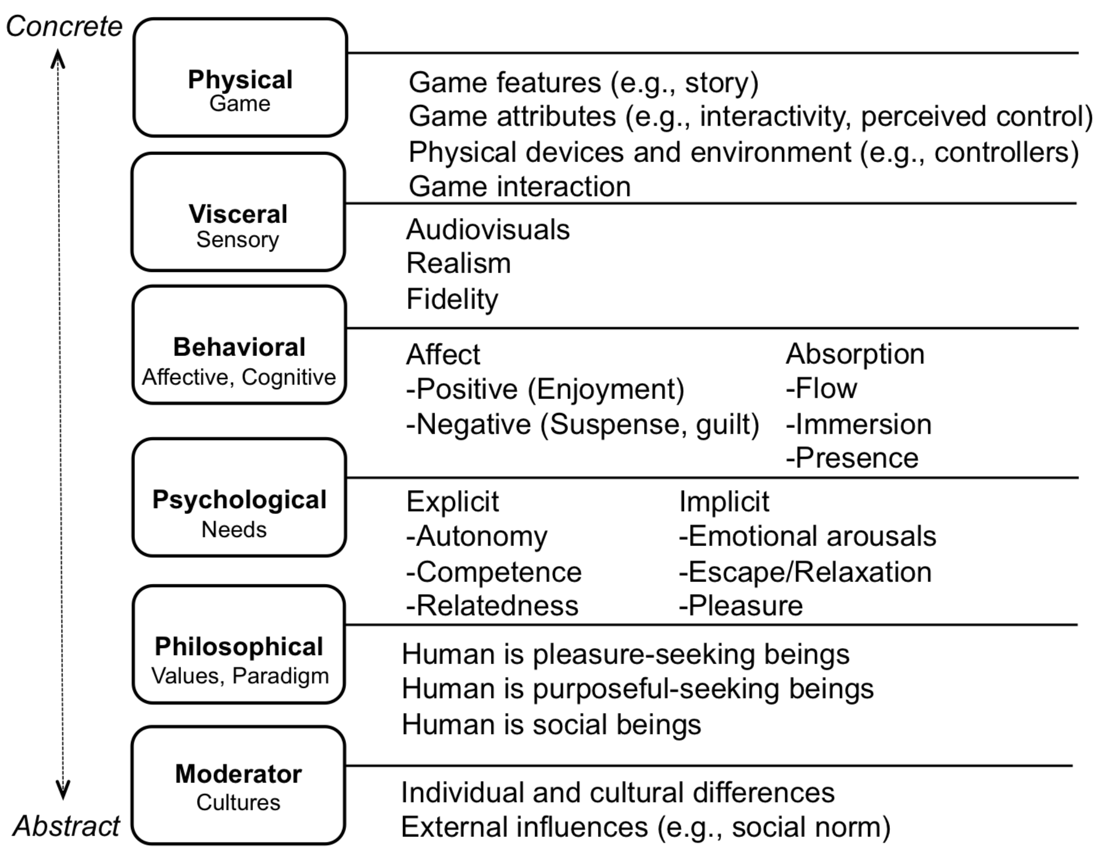


Figure 3.1: Six-Dimensional Engagement Framework.

3.7.2 Relationships between Engagement and Other Related Concepts

The relationships between engagement and other concepts including needs satisfaction, emotions, enjoyment, flow, immersion, presence may provide information for the future development of engagement model or framework.

Needs satisfaction and emotional arousals are often viewed as the main reasons for game engagement. The ability of games to satisfy psychological needs and to elicit various strong emotions is often regarded as a key explanation for why people play games. The concept of dual processing models [46] may partly explain the coexistence of needs satisfaction and emotional arousal. Dual processing models refer to the idea that people tend to process information in one of two ways - intuitively (or implicitly or unconsciously) or explicitly (or consciously). This generally argues that when responding to a stimulus, we can process it via careful and deliberate consideration, or via simply heuristic or peripheral cues. One may engage games with simple processing and it is experienced as enjoyment (emotional arousal). In contrast, some may engage games with deeper processing and experience the games more akin to appreciation (needs satisfaction).

Enjoyment and absorption (flow, immersion, and presence) are often viewed as the outcome of a good gaming experience. They may also be seen as players motives (e.g., player play games so to experience enjoyment and/or absorption/escape). If we were to consider the cognitive and affective axis, enjoyment can be regarded as an affective outcome of a great gaming experience, while absorption describes the outcome on the cognitive axis.

Measuring engagement should take account for the reasons/motives as well as the various outcome of gaming experience. Researchers have taken multiple approaches to measure engagement, primarily questionnaires from the dimension of needs satisfaction [102], absorption [17], immersion [58], affect and challenge [54]. Among all the questionnaires, enjoyment is the most highly asked [76]. There are tendency that game engagement to be treated synonymously or used interchangeably with enjoyment, flow or other related concept but limiting the study of engagement to only one would fail to provide a more comprehensive account of other determinants of engagement including player motives and game features.

Chapter 4

STUDY 2: LARGE-SCALE ONLINE STUDY

From previous theoretical study, this study aims to identify practical game features that makes game engaging, which is not much explored in Study 1. To explore game engagement from gamers perspective, we conducted a Web survey to examine the motivation of gamers. We choose to conduct a Web survey because of its accessibility to large groups of diverse populations. The Web survey was advertised on high-traffic game forums, which targets diverse gamers who regularly play games. We used open-ended questions to identify principles or factors that motivate gamers. Additional questions included demographic information.

To investigate the motivation behind gamers, our web survey focused on two aspects: (i) engagement and (ii) disengagement of gamers. Both of these aspects of gameplay affect how gamers choose to play and engage in games and therefore provide a useful base from which to begin exploring game engagement. To examine these two aspects, we asked the following open-ended questions to the respondents (see Table 4.1):

Q1: Please list your favorite game and share with us why you like the game.
Q2: Are there any games where you just annot stop playing? If yes, can you share why?
Q3: In your opinion, what are the most important elements for a game?
Q4: Can you share with us what makes you like play games?
Q5: What is usually the reason you stopped playing a certain game?

Table 4.1: Web survey questions

4.1 Methodology

4.1.1 Participants

A total of 203 respondents participated in this web survey, ranging from 13 to 48 years old with a mean of 29.18 years. With respect to geographic nationality, 32% of respondents were from Thailand, 29% from United States, 9% from United Kingdom, 7% from Canada, 3% from Japan, 2% from Australia, 2% from China, 2% from Ireland, while others accounted for 14%. 6% of the respondents were female and 93% were male. With regard to gamer experience, 52% of survey respondents reported playing games every day; 26% play 4-5 times a week, 10% play twice a week, 9% play once a week, and 2% reported they never play games. The respondents have the most game genre-related interests with First Person Shooters (71%), Role playing (71%), Turn-based strategy (57%), Real-time strategy (48%), Multiplayer (42%), Fighting and Battle (34%), Racing (32%), Sims and Virtual life (25%), Music (25%), Sports (22%), Action and Action Adventure (15%), Multiplayer Online Battle Arenas (15%), and Puzzles (8%).

4.1.2 Procedure and Apparatus

We prepared our web survey via Google Form and the survey was advertised on high-traffic game forums. Before beginning the survey, respondents were told the general aim of the study. Then respondents were requested to fill in demographic information. Then before the actual survey questions, respondents were asked whether they play games. Respondents who do not play games were redirected to the end of the survey. Otherwise, respondents were directed to all survey questions. This initial choice was to ensure that results were only from gamers.

4.2 Results and analysis

To identify patterns from the survey, frequency data were collected and measured. We used sentiment analysis approach using IBM SPSS Text Analytics for Surveys to identify conceptual themes. The SPSS software extracted keywords from respondents answers, and allowed us to classify the themes by grouping synonymous terms with similar semantics (e.g., friends, team, community -> "Peer motivation"; role-playing, narratives, characters > "Story"), and to measure the frequency data by number of occurrences reported from unique participants. Figure 5 shows the emerging themes and their frequency for each question.

	Story	**Challenge**	**Peer motivation**	**Exploration**	**Aesthetics**
Q1	27%	19%	16%	11%	7%
Q2	9%	7%	10%	5%	0%
Q3	70%	35%	23%	33%	34%
Q4	62%	47%	13%	61%	36%
Q5	31%	48%	21%	48%	12%

Table 4.2: Frequency data of emerging themes from the web survey.

The results revealed five major themes as follows: (i) Story, (ii) Challenge, (iii) Peer motivation, (iv) Exploration, and (v) Aesthetics. We described each theme with sample answers from the respondents.

Story

Story promotes a strong sense of emotional involvement by bringing purpose and relevance to the goal of the game. Story is often mentioned by respondents together with role-playing, difficult choices-making, and multiple story paths which make the overall game, engaging. For examples,

"The narrative, story, and choices that Mass Effect 2 allows you to make all impact on are why its my favorite game. The interactions with the characters of your crew give you a direct link of emotion to them and want to see them all survive." [P1, Q1]

"Gaming is the ability to be someone else and do things that you normally cannot do. I believe that anyone who likes to play games shares this view and is a large reason as to why they enjoy them." [P54, Q4]

"Story brings purpose to the game. Its the whole reason that drives me. I usually dont play a game long if without a good story." [P54, Q3]

One interesting question raised to us is: If gamers enjoy story, why do they not prefer movies or novels to games? Our web survey revealed that it is due to the amount of interactivity and impact on the overall story that makes the games more preferable for gamers. For example,

"Games beat TV or movies. It is so great that you can actually impact the story, interact with the characters and see them grow. Generally, I just prefer interactive entertainment to just watching movies or TV." [P105, Q4]

Challenge

Respondents reported that achieving challenges provides great sense of accomplishment and mastery. Most respondents prefer challenges that require practice and creativity, rather than challenges that based on luck or currency. This is possibly because the sense of accomplishment is greatest when players skills are fully utilized in achieving a challenge.

"This game (Dark Souls) is extremely difficult. It is one of the hardest games ever made. However, when you finally beat a part that you've been stuck on it feels so satisfying. The feeling of defeating your massive foes is unmatched." [P164, Q1]

"It requires skill and patience to make progress in these games but heavily punishes you if you let your guard down. I get a real sense of achievement when I make progress in these games." [P91, Q3]

The ability to retry after failure appears an important mechanism that sustains players interest and allows players to consistently hone their skills for increasingly difficult challenge.

"I always find a challenge playing Diablo 3, its learning curve is similar to Dark Souls; in which, each time you die (because you will) you learn a little bit more about strategy, spacing, and skill rotation, then you apply it." [P3, Q1]

Respondents prefer balanced combat system with suitable challenges to their skill level. Otherwise, they are likely to withdraw.

"However, I can enjoy the story line of the games only when the game is not too hard and

not too easy. When I achieve a goal (quest), I'm motivated to play more. If I can achieve another goal before my motivation runs out, it becomes addictive. When my motivation runs out, I find a new game." [P99, Q5]

"Warhawk: The perfect balance between all classes, and nearly everything." [P143, Q1]

"when I put extra time and effort into something, I generally should stand out. In this case, different looking gear to show the work I put into it. Yes it may be sad to get gratification from a game, but if I put extra time and effort into a game and have the opportunity to stand out taken away, there is no reason to play." [P68, Q5]

Peer motivation

Peer motivation can enhance and prolong players engagement by adding social experience and quality competition or cooperation. Some examples include:

"in using all of this (our skills and strategy) to outwit, and outplay real people who are trying to do exactly the same thing is so satisfying. Single player lacks this competition, as we are far from creating smart, computer controlled bots." [P6, Q3]

"When I play this game with a few friends, there is so much strategy involved and how we plan out our attacks that it takes us almost 2 1/2 hours to do one battle. I just got so addicted." [P7, Q2]

"I liked social games because I got to know many people, and to a place where I feel valued and belonged." [P172, Q2]

However, on the other hand, it seems overly demanding social activities may also degrade players engagement:

"If it becomes too demanding like you have to log in everyday or your guildmates will kick you out I will start to become bored and might stop playing eventually." [P3, Q5]

Exploration

Exploration refers to players freedom to explore a virtual world filled with different possibilities or refers the players ability to explore different solutions to a challenge.

"The fun factor largely determined by what I am allowed to do. For example, if I get tired

of playing story missions in Saints Row the Third, I always have the option of performing tasks, and virtually any action in the game contributes to the leveling system. Each new level allows new perks, which is a game mechanic in itself." [P203, Q3]

"I really liked chess kind of games, even in digital form. Chess is so fun because everyime you play, its different. You have thousands of way to explore and win the game. There's freedom to experiment and it's all my strategic choices." [P185, Q1]

When a game lacks exploratory momentum, players tend to withdraw the games over time. For examples,

"When I feel there is nothing ahead or become very linear of the game, I simply quit the game." [P90, Q5]

"It just gets too repetitive and I can't stand to play anymore." [P25, Q5]

Aesthetics

Aesthetics, which largely refers to the "look and feel" of the game, is largely determined by the games graphics, audio, and realism. While most respondents do not consider aesthetics as primary, they also feel that aesthetics does contribute to the overall enjoyment of a game. In that sense, aesthetics appears equally important.

"I find that graphics and sound certainly make a game better, but only by improving what's already there (the icing on the cake). A game can look and sound good, but if the gameplay experience is bad, I'd rather watch a movie." [P73, Q3]

"Story is most important, but I appreciate games for being multifaceted: visuals, sound, music. It is more immersive." [P166, Q3]

4.3 Discussion

We have investigated game engagement from gamers perspective and identified five pertinent concepts. We followed up by updating our initial framework to include the five game features (see texts in red in Figure 4.1).

This study improves our practical understanding regarding game engagement. In the next studies, we conducted multiple evaluation studies focusing on different dimensions of

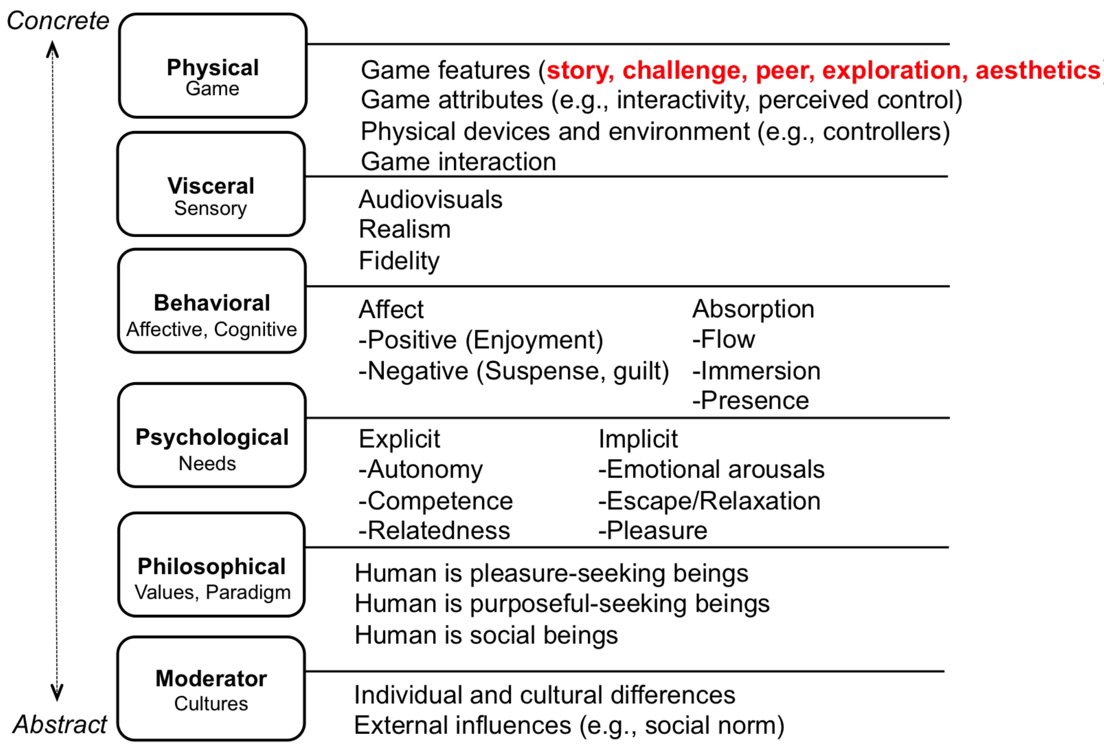


Figure 4.1: Updated Six-Dimensional Engagement Framework

the framework using gamification and full-body games as case studies.

Chapter 5
STUDY 3: EMPIRICAL STUDY IN GAMIFICATION (I)

This study aims to evaluate the effect of game engagement using gamification as a case study.

5.1 Methodology

5.1.1 Game elements under testing

We selected challenge from our framework as a initial point of investigation (see 5.1). Challenge was designed under three principles: goal, progress and competition. Goal was implemented in form of badge as a technique to inform users the goal. Progress was implemented in form of points and progressbar. Competition was implemented using challenge and leaderboard.

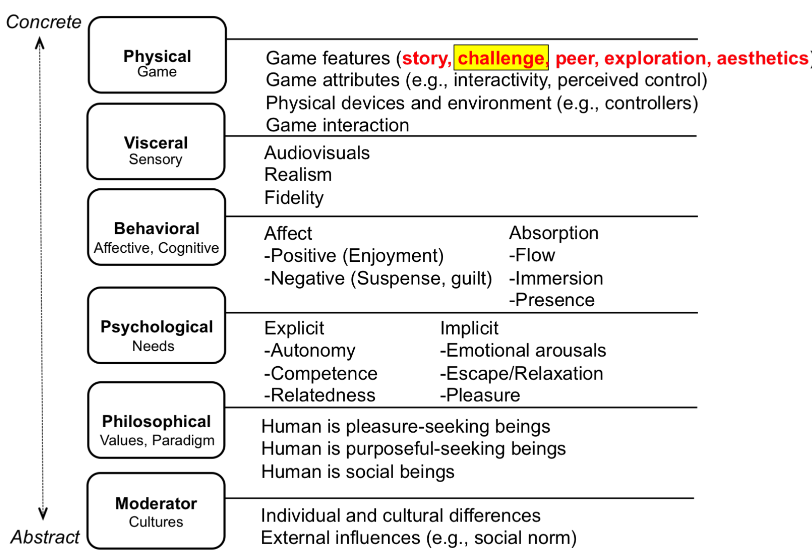


Figure 5.1: Study 3: Evaluating challenge

The system provides lessons and challenges for students to complete. The system establishes a reinforcement schedule where badges will be awarded when students achieve certain milestones. Game objectives, progress-bar and points are consistently presented on the top part of the interface to indicate progress. Students can check their ranking in the leaderboard at any time. See Figure 5.2 for screenshot of the interface.

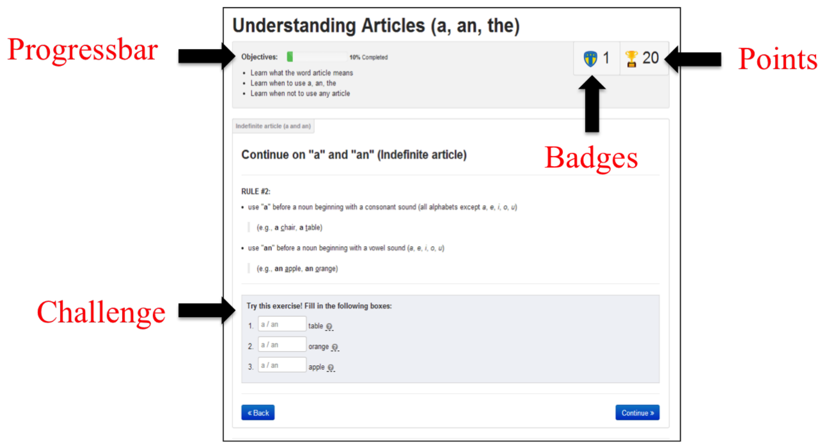


Figure 5.2: Screenshot of interface used in Study 3

5.1.2 *Design*

This study used a between-subject design. Two systems (no game, with game) were tested. With-game (WG) group contained the aforementioned game elements.

The experimental task was about learning English usage of articles. Prior to the experiment, participants were asked to complete a prettest measuring their level of competency about English articles. After the experiment, a posttest with similar questions was issued to measure learning gain.

5.1.3 *Participants*

19 university students (8 females, 11 males) were recruited. The age mean was 21.78 years old. Student frequency with games was 2% (>20hrs a week), 5% (11-20hrs a week), 8%

(8-11hrs a week), 12% (5-8hrs a week), 21% (3-5hrs a week), 31% (1-3hrs a week), and 21% (I don't play games).

9 participants used NG and 10 participants used WG. Competency of English articles was balanced between the two groups.

5.1.4 Apparatus

An Intel core i7-2600 3.40GHZ PC with 8GB Ram and Window 7 Enterprise was used for the experiment. The systems were implemented using web technologies (i.e., HTML, Javascript) and were installed on a local network to minimize any possible network or speed problems. Pretest and Posttest were prepared by English specialists in paper-based form.

5.1.5 Measurement

We used learning gain as first measurement of engagement. We also used a 7-likert scale questionnaire measuring user satisfaction after using the system. Semi-structured interviews were employed to further assess user preferences.

Figure 5.3 shows the engagment model of this study.

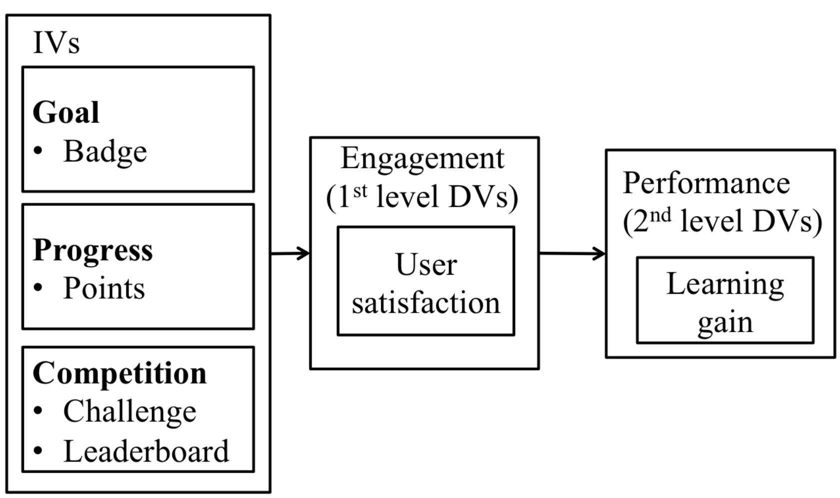


Figure 5.3: Study 3's engagement model

5.2 *Results and analysis*

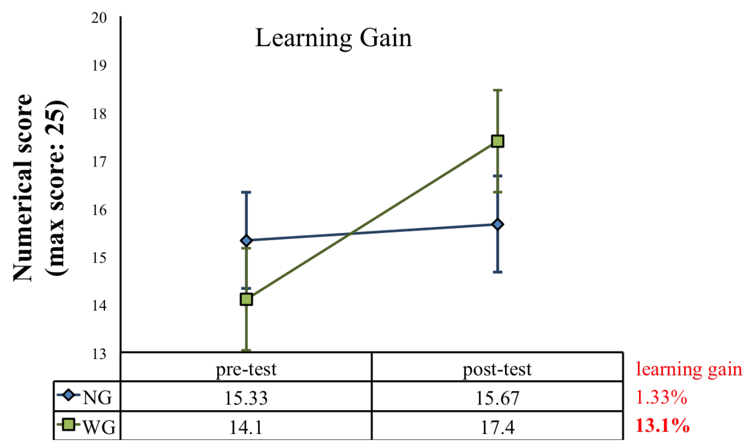


Figure 5.4: Learning gain: NG vs WG (p<0.001)

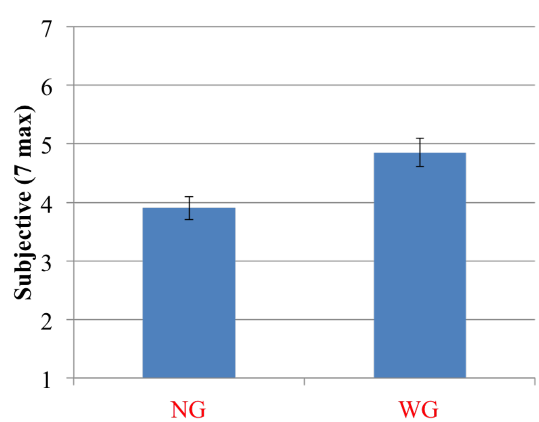


Figure 5.5: User satisfaction: NG vs WG (p<0.01)

Learning gain and user satisfaction were analyzed using ANOVA test. The learning gain in WG was significantly higher than NG ($F_{1,17}$=13.67, $p<0.001$). User satisfaction in WG was also significantly higher than NG ($F_{1,17}$=11.66, $p<0.01$). Figures 5.4 and 5.5 show the results.

Our semi-structured interviews revealed positive feedback of students in using the WG system. They also suggested adding other game elements such as Story to the systems.

5.3 Discussion

This study served as an initial study to evaluate the effect of game engagement in design. Results showed that game engagement can help improve user performance (learning gain) and user engagement (user satisfaction).

The limitation of this study is about the few game elements under testing. In addition, engagement was only measured using one scale, i.e., user satisfaction. Our next study aimed to address these two gaps.

Chapter 6

STUDY 4: EMPIRICAL STUDY IN GAMIFICATION (II)

This study aims to further evaluate the effect of game engagement by adding game element under testing and to more systematically evaluate engagement (see yellow-highligted parts in Figure 6.1).

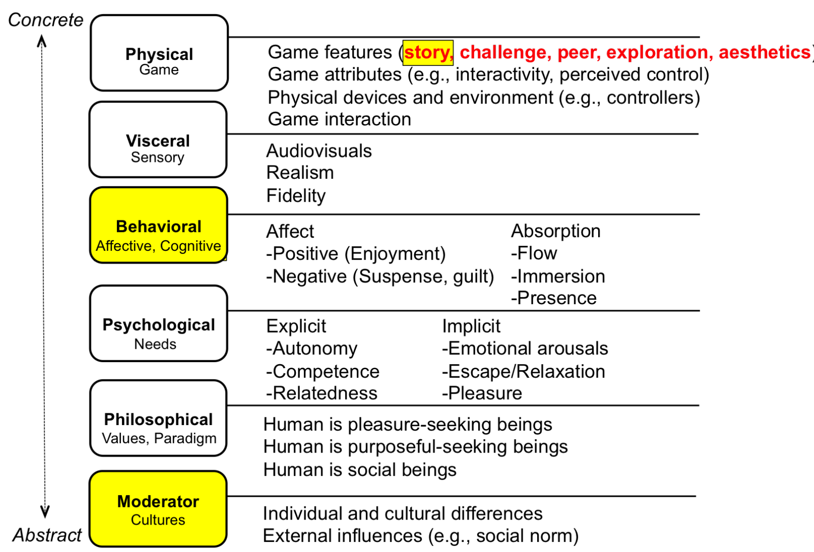


Figure 6.1: Study 4: Evaluating story and behavioral and moderators dimensions of the framework.

6.1 *Methodology*

6.1.1 Game elements under testing

Aside from challenge, we selected Story from our framework for investigation. Story was designed with four components: pictures, narratives, choices and story point.

Users role-play through the narratives, at the same time, making choices that will have impact on the story ending. There are two types of choices: one that is related to the learning materials itself (e.g., you found ? orange, what is the correct English article?), and one that is not related to (e.g., which path will you choose? the forest or the roadway?). Our rationale is to hide the learning within the game objectives so to make students forget that they are learning.

To make the choices meaningful, each time students failed to make the correct choices costs 1 story point which will impact the story ending.

To ensure an engaging game story and settings, we developed this game based on the story section of a popular video game "Zelda: Lost Woods". We used high-quality cartoonic pictures to inspire the imagination while students are going through the game story.

See Figure 6.2 for screenshot of the interface.

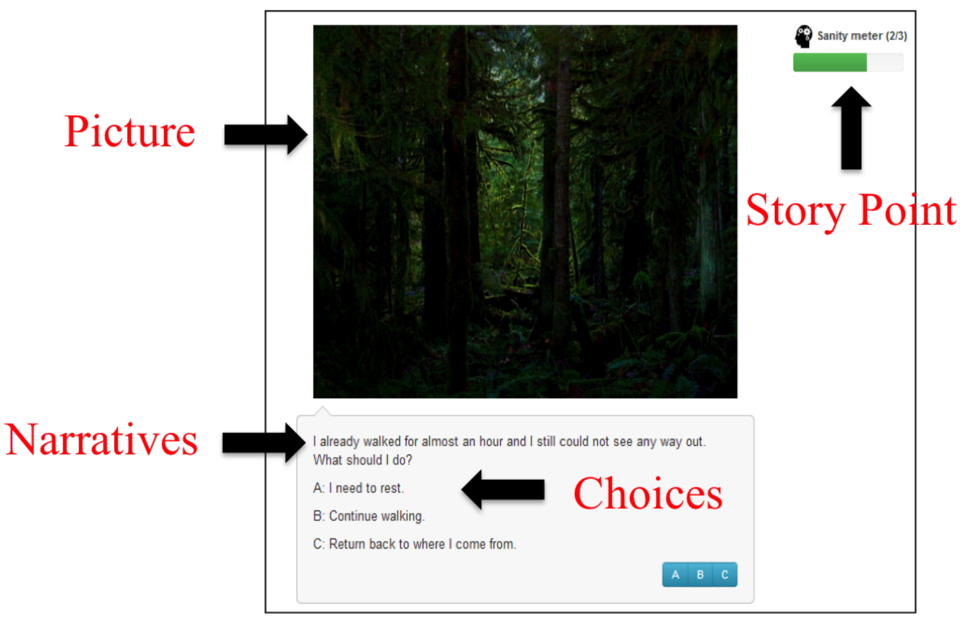


Figure 6.2: Screenshot of interface used in Study 4

6.1.2 Design

This study used a within-subject design. Three systems (FunEnglish, LostWoods, LostWoods+) were tested. FunEnglish (FE) is the challenge system used in Study 3. LostWoods (LW) is the story system mentioned in previous subsection. LostWoods Plus (LW+) extend LostWoods to include progressbar, badges, points and leaderboard. By introducing LW+, we aimed to explore whether the addition of progressbar, badges, points and leaderboard has any positive or negative effect on LW.

The experimental task was same as Study 3.

6.1.3 Participants

26 university students (14 females, 12 males) were recruited. The age mean was 21.04 years old. Student frequency with games was 8% (>20hrs a week), 4% (11-20hrs a week), 0% (8-11hrs a week), 12% (5-8hrs a week), 31% (3-5hrs a week), 27% (1-3hrs a week), and 19% (I don't play games).

6.1.4 Apparatus

An Intel core i7-2600 3.40GHZ PC with 8GB Ram and Window 7 Enterprise was used for the experiment. The systems were implemented using web technologies (i.e., HTML, Javascript) and were installed on a local network to minimize any possible network or speed problems.

6.1.5 Procedure

In the experiment, participants were tasked with playing the three systems. First, all participants were informed about the aim and the general procedure of the study. Then participants were asked to fill in demographic info regarding their gaming frequency. Then participants were asked to play the three games (FE, LW, LW+) in a counterbalanced order. They were given a 5-minute break between each session of games. A questionnaire session and semi-structured interview were conducted afterward.

6.1.6 Measurement

Based on our framework, we measured engagement in two dimensions: cognitive and affective. Cognitive engagement relates to attention and effort. Affective engagement relates to enjoyment.

For cognitive engagement, four subscales were measured - behavioral engagement (BE, 6 items, e.g., I pay attention), behavioral disaffection (BD, 6 items, e.g., I do just enough to get by), cognitive strategy use (CSU, 13 items, e.g., When I study for a test, I try to put together the information.), and self-regulation (SR, 9 items, e.g., I ask myself question to make sure I know the material that I have been studying).

For affective engagement, two subscales were measured - emotional engagement (EE, 6 items, e.g., The activity is fun) and emotion disaffection (ED, 6 items, e.g., I feel bored).

We adapted these scales from EsVD [108] (Engagement versus Disaffection with Learning) questionnaire and MSLQ [90] (Motivated Strategies for Learning Questionnaire).

Thus a total of 46 items in 6 engagement subscales were measured. All items followed a 7-point Likert scale (1 as not at all true of me and 7 as very true of me).

For collecting qualitative data, we conducted semi-structured interviews at the final stage: What do you like and dislike about each game? If you want to learn something using games, what kind of gameplay would you prefer?

Figure 6.3 shows the engagment model of this study. Yellow-highlighted objects show the updated part of Study 4 from Study 3.

6.2 Results and analysis

Engagement scores of the three systems and the effect of gamers vs. non-gamers were analyzed using ANOVA tests and Posthoc comparison with Bonferroni correction.

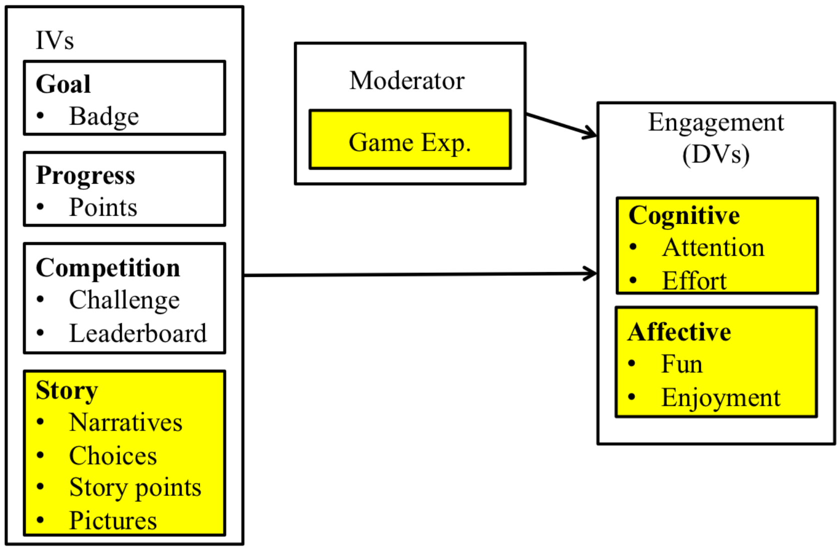


Figure 6.3: Study 4's engagement model

6.2.1 FE vs. LW vs. LW+

Repeated measures ANOVA shows a significant difference ($F_{2,75}$=4.68, $p<0.01$) in EE between three systems. A post-hoc comparison confirms a difference ($p<0.01$) in EE between FE (M=4.82, SD=1.13) and LW (M=5.86, SD=0.92). Figure 6.4 shows that LW has the highest rating in BE and EE, and lowest rating in BD and ED.

To our surprise, LW+ failed to show any significant improvement over LW. Conversely, LW+ scores generally lower than LW. In the interview, participants commented that FE became repetitive over time, that LW was enjoyable and immersive, and that LW+ was also enjoyable but composed of unnecessary, distracting components on the interface. Table 6.1 summarizes participants' comments.

6.2.2 Gamers vs. non-gamers

We classified our participants into three gamer categories: non-gamers (0-3 hours a week), moderate-gamers (3-11 hours a week), and frequent-gamers (>11 hours a week). Repeated measures ANOVA shows a significant difference ($F_{2,69}$=3.35, $p<0.05$) between the three cat-

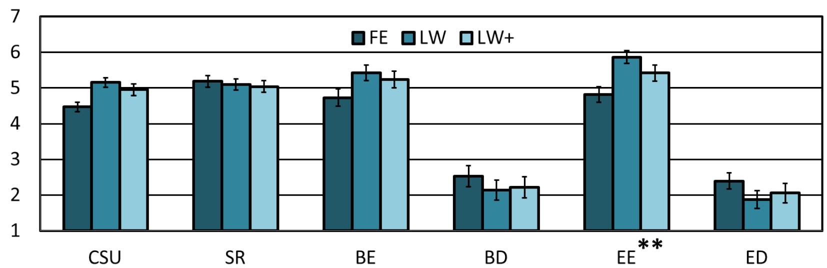


Figure 6.4: Engagement scores of the three systems. Emotional engagement (EE) shows a significant difference (p<0.01) between FE and LW. (Error bar show 95% CI)

	Positive comments	**Negative comments**
FE	• feel confident with the achievements • compare rankings with friends • simple to understand • immediate feedback • interactive • increasing difficulty • clear purpose	• repetitive • unmeaningful achievements • lack of graphics and sound effects
LW	• story is fun • multiple endings • choices • immediate feedback • interactive • immersive	• more choices • should have cut-scenes • should have better music • too short gameplay • too stressful with the choices
LW+	• story is fun • multiple endings • choices • interactive • immediate feedback	• points and badges do not add much • progressbar disrupts immersion • too many things on the interface

Table 6.1: Positive and negative comments between educational games.

egories of gamers in the subscale of ED. A post-hoc comparison with Bonferroni correction confirms the difference (p<0.05) in ED between non-gamers and frequent-gamers. No significant effect of gaming frequency on other engagement subscales was found. Figure 6.5 shows that frequent-gamers have the highest level of BE and EE. However, frequent-gamers also have the highest level of BD and ED.

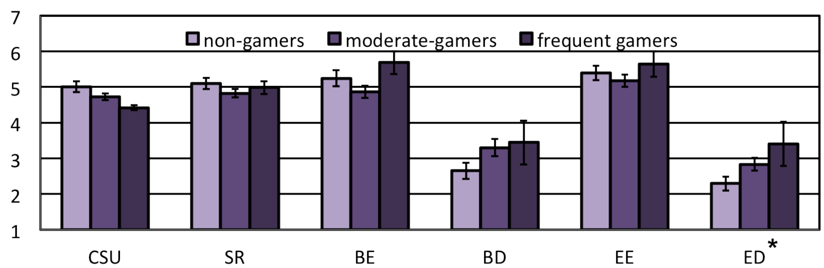


Figure 6.5: Engagement vs. gaming frequency. Significant difference in ED (p<0.05) between non-gamers and frequent-gamers was found.

Non-gamers	**Frequent-gamers**
• simple interfaces	• engaging story/scenario
• effective for learning	• graphically appealing
• takes small amount of time to play	• fun gameplay
• playing with friends	• rich interfaces
• humor	• high amount of interactivity
• graphically appealing	• audibly appealing
	• effective for learning
	• playing with friends

Table 6.2: Preferences between non-gamers and frequent gamers (order by descending frequencies).

Frequent-gamers achieved generally higher levels of engagement. However, to our surprise, frequent-gamers also achieved higher level of disengagement. Our interview found that students who regularly play games tend to have higher expectation, negatively commenting on the need to include additional features such as cut-scenes, animations, or sound effects. Conversely, students who do not regularly play games tend to enjoy our games, complimenting our games as having simple interfaces and taking a small amount of time to play. Table 6.2 summarizes differences in preferences between non-gamers and frequent-gamers.

6.3 Discussion

This study served as a second study to evaluate the effect of game engagement. Results showed positive effect on user engagement. Results also confirmed the differences between gamers and non-gamers.

Although LW appears to be superior over FE, we found that users who are already highly motivated to learn English often prefer FE, commenting that it is more convenient for learning. This leaves room for further investigation on two types of users: high-motivated user and low-motivated user.

The major limitation of this study is about the generalizability of the results. That is, the effect may be specific to the experimental tasks and thus may not be generalizable to other learning tasks. Our next study aimed to address this gap.

Chapter 7

STUDY 5: EMPIRICAL STUDY IN GAMIFICATION (III)

This study aims to further evaluate the effect of game engagement by using abstract tasks. Using abstract tasks can remove possible confounding factors introduced by the use of specific learning applications. This approach was inspired by the *Reductionism* theory [93], a philosophical research approach that attempts to understand a complex set of phenomena through its simpler, fundamental elements. Specifically, instead of teaching a particular subject of study, we tested how well people can master and improve some of the basic cognitive capabilities required by the act of learning in various disciplines, such as the memory and problem-solving skills, using abstract tasks. This allows us to investigate the effect of game elements on particular aspects of learning at a micro-level (e.g., game element A, B, C can improve the cognitive skill X, but not Y). Furthermore, this approach helps minimize possible complications introduced by the quality of the materials or the user's personal interest/skills, and thus increases the overall generalizability of the results.

7.1 Methodology

7.1.1 Selecting Learning Tasks

We adapted a number of standard fundamental learning tasks from cognitive psychology for our study. We had three criteria for task selection - (1) the learning tasks should target some fundamental cognitive skills required in learning various disciplines; (2) the learning tasks should be commonly used and should be empirically validated with satisfactory psychometric properties; (3) the number of learning tasks should accommodate a reasonable experimental timeframe and should not overwhelm users with excessive cognitive overload. Given these considerations, we selected *memory* and *problem-solving* for our investigation, since they are

	Learning Aspect	Sub-aspects	Measures
Yes-No Recognition	Memory	Recognition	Total number of correct responses
Cued-Recall	Memory	Recall	Total number of correct responses
Free-Recall	Memory	Heavy Recall	Total number of recalled items
Wisconsin Card Sorting	Problem-Solving	Flexibility	Total number of correct matches
Tower of London	Problem-Solving	Planning	Average moves per problem, total time spent
Wason Selection	Problem-Solving	Deductive Reasoning	Number of failed attempts before getting right

Table 7.1: Overview of selected learning tasks.

two of the essential prerequisites for effective learning in many disciplines such as linguistics, computer science and engineering. In other words, in the scope of this paper, we specifically refer to "learning" as the increase in user's memory and problem-solving capabilities through repeated practice-based training.

See Table 7.1 for the complete list of abstract tasks.

Memory

In psychology, there are two types of memory retrieval: recognition and recall [94]. We selected the following learning tasks for training these two functions of declarative memory respectively (Table 7.1):

- Yes-No Recognition Task (YNR) [94, 110] (Recognition)
- Free-Recall Task (FR) [14] (Recall)
- Cued-Recall Task (CR) [14] (Recall)

Yes-No Recognition Task (YNR). Participants are first exposed to a list of 15 items, which they are given 30 seconds to memorize. These items are randomly drawn from a pool of 150 candidate words and numbers that frequently occur in linguistics, science, and engineering related fields. After the 30 seconds exposure, participants need to answer a set of 25 yes/no questions on whether a certain item appears in the list they have just seen.

Free-Recall Task (FR). Similar to the previous task, we gave participants two minutes to remember a list of 15 randomized items. Then participants are asked to recollect as many items as possible in any order within three minutes.

Cued-Recall task (CR). Participants first view a list of 10 pairs of randomized items for two minutes. Then, given a set of 10 cues, participants are asked to recall the item which was originally paired with the cue.

Each of these memory tasks takes around five minutes. We measure the total number of correct responses as an indicator of performance.

Problem-solving

We selected three most commonly-used and empirically validated learning tasks with different execution functions of problem solving, including flexibility, planning, and deductive reasoning (Table 7.1):

- Wisconsin Card Sorting Task (WCST) [45] (Flexibility)
- Tower of London Task (TOL) [116] (Planning)
- Wason Selection Task (WS) [118] (Deductive Reasoning)

Wisconsin Card Sorting Task (WCST). Participants are shown four cards that each contain a set of geometric designs varying in form, color, and number of elements. Then participants are asked to pick out a fifth card that matches any of the four cards based on either form, color, or element number, but are not told the rule of matching in advance. However, participants received feedback on whether each assignment is right or wrong. The rule of matching randomly changes during the course of the task. In our implementation, the matching rule changes randomly after every five consecutive correct responses, but participants are not informed of this change. The task terminates upon the completion of 64 rounds of card matching or when time runs out after three minutes, whichever comes first. WCST was commonly used to measure participants' flexibility in problem-solving, as the task requires them to dynamically adapt to the changing rule. Performance is based on the total number of correct matches.

Tower of London Task (TOL). Participants are shown six colored disks on three pegs and are asked to shuffle the disks, one move at a time, to make a given random arrangement. The task terminates after solving five such problems. This task was used to measure participant's problem-solving and planning capability. The average number of moves per problem and the total time spent on solving all five problems are indicators of participants' planning ability.

Wason Selection Task (WS). Participants are given four cards and a statement regarding the content written on both sides of the cards. Participants need to point out which cards they want to flip over to see the back side in order to determine whether the statement is true or not. For example [119], four cards are labeled “A”, “B”, “4”, and “7” respectively on the front side, and the statement requiring judgement says “*If a card has a vowel on one side, then it has an even number on the other side*”. Participants have to decide which cards to turn over so that they can determine the validity of the statement within a minimum number of steps. The only correct response in this example is “A” and “7” [119]. Each task only consists of one such problem. This task reflects participants' deductive reasoning skills. We measure task performance using the number of failed attempts before getting the problem solved. This task takes around two minutes on average.

7.1.2 Developing Tools for Testing

We devised two interactive online learning environments (control, experimental) for the same sets of abstract tasks. The control version was designed without any game elements. The experimental version was designed using challenge, where challenge was designed under the three same principles as Study 3 - goal, progress and competition. Goal was implemented in form of badge as a technique to inform users the goal. Progress was implemented in form of points. Competition was implemented using leaderboard.

Points take the form of scores provided at the end of each task, so as to provide participants with a sense of progress toward mastery.

In contrast, badges are awarded only upon the completion of certain challenges. We designed a total of 42 badges, seven per learning task. Each badge contains a visual icon, a name, and an instruction of the corresponding unlocking conditions (example in Figure 7.1). Participants receive a notification each time they successfully unlock a certain badge. We included badges in our design because they can help set up clear goals and challenges for participants to strive for, and create a sense of achievement and immediate feedback once granted. In addition, the badges available to the participants should reflect the spectrum of task difficulty, and should be visible to users. The thresholds of each badge level (e.g., Figure 7.1) were determined through playtesting with nine users prior to the actual study.

The Leaderboard displays the normalized, accumulated task performance scores of all participants in descending order. Participants can further filter and sort the scores according to each abstract task. The Leaderboard allows participants to compare their performance with others, motivating participants by stimulating an indirect competition. Since social presence of others can possibly impact the effectiveness of the leaderboard, we asked several users to play the experimental version to populate the leaderboard with data, so that our initial set of participants can start comparing their relative performance with others.

Since we do not want to overburden participants with extra efforts to jump to another page to view the points, badges, and leaderboard, we embedded these game elements in the task interface, as shown in Figure 7.2.

7.1.3 Design

We split our study into two sessions, each focusing on one targeted cognitive skill of learning, i.e., memory or problem-solving. We counterbalanced the order across participants. We required participants to complete three trials for each learning task, but allowed them to conduct more trials if they wanted to. For each participant, the two sessions were held on the same day, with a 20-minute break between. Each session took about 60 minutes. All

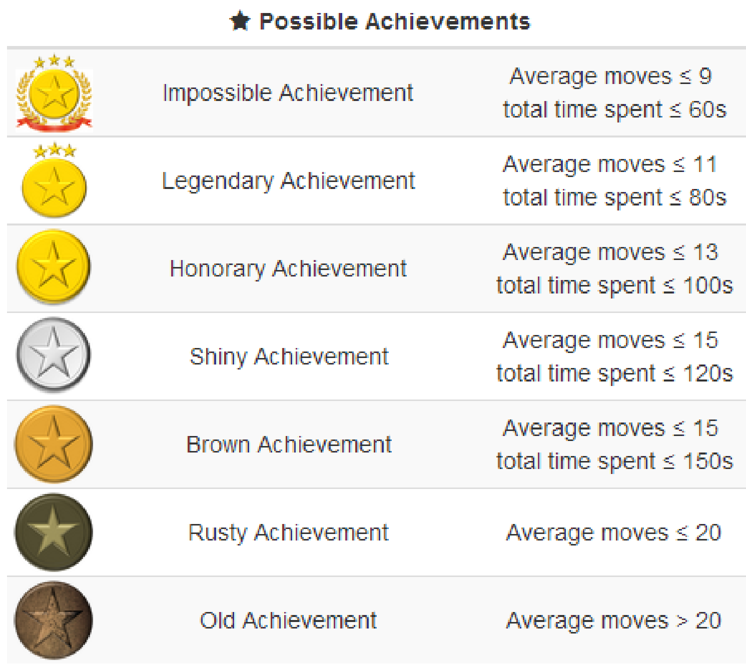


Figure 7.1: Screenshot of badges used in our study in Tower of London task.

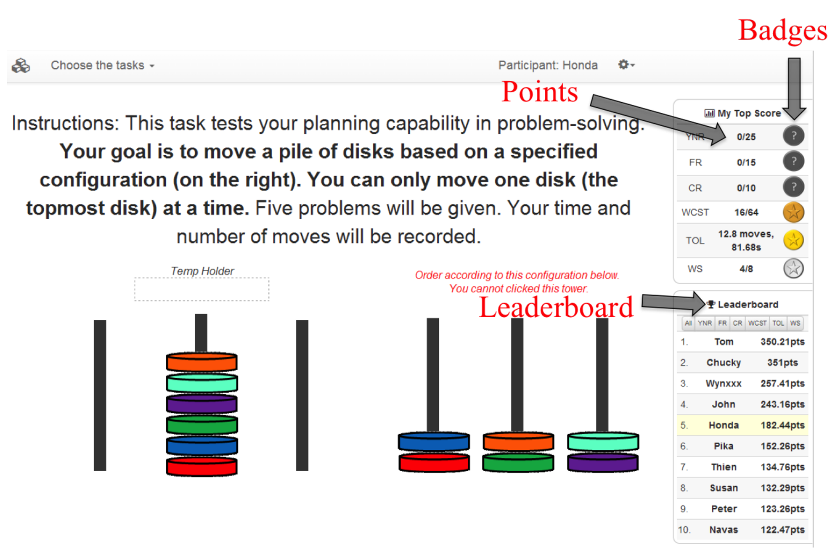


Figure 7.2: Screenshot of experimental version of interface during the Tower of London task.

experiments were conducted in the afternoon to ensure that individual performance was not affected by the time of the day.

Prior to the study, each participant took a cognitive test using the six abstract tasks. The cognitive test aimed at evaluating each participant's memory and problem solving skills to sort participants into two balanced groups.

When finishing all tasks, participants completed a subjective assessment questionnaire measuring their general attitudes and preferences.

7.1.4 Participants

To test our hypotheses, we conducted a between-subject study. Two groups of 15 participants (age 19 to 25, M=22.1 years, 4 females) were assigned to one of the two systems - control and experimental. We kept cognitive performance, age, and gender balanced between the two groups. Participants were recruited through the university announcement portal. All participants had experience with video games with a mean frequency of two to five hours per week. None had experienced any of the learning tasks. Each was paid $10.

7.1.5 Apparatus

The study was conducted on a high-performance desktop computer (Windows 7) connected to a Tobii X120 eye-tracker. The data from the eye-tracking device was recorded and transformed into a heatmap visualization and fixation percentage using Tobii Analytics SDK. The learning system was implemented using JavaScript and HTML5 and was run on a Chrome browser.

7.1.6 Measurement

In this study, we introduced three measures to assess the effect of the game elements. The first measure was task performance (summarized in Table 7.1). The second measure was the subjective feedback collected through a post-study questionnaire, which was correlated with each participant's gaze behavior on the task interface (measured by a Tobii eye-tracking device throughout the study). The last measure was the number of extra trials that participants performed in addition to the three compulsory trials. Specifically, we hypothesized that the experimental group will significantly outperform the control group in all three measures.

7.1.7 Experimental Protocol

A few days before the experiment, we asked all 30 participants to fill out a general demographic questionnaire, providing information on their age, gender and prior gaming experiences. We also obtained their scores in the cognitive tests to determine their memory and problem-solving capabilities. Based on these results, we balanced participant assignment between the experimental group and the control group.

It is important to note that the group assignment in the memory session is independent of the group assignment in the problem-solving session. In other words, a participant belonging to the gamified group in the memory session, may belong to the non-gamified group in the problem-solving session. This is because an individual's memory and problem-solving ability can be very different, and thus it is difficult to take both capabilities into account at the same time when trying to balance the pretest performance. Although this arrangement may yield possible carry-over effects from the first session to the second session, we tried to minimize the influence by counterbalancing the order of sessions across all participants.

Both sessions followed similar protocols. We present the protocol of the memory session as an example:

1. Training in Learning Tasks (5 minutes): Participants got familiarized with the three memory learning tasks through practicing some example questions.

2. Training in System (5 minutes): In the experimental group, experimenters demonstrated how the points, badges, and leaderboard worked. Participants were allowed to try the system after receiving instructions. In the control group, participants were shown only the task interface.

3. Learning Task (45 minutes): Participants then proceeded to the three memory learning tasks, each task containing 3 trials. The order of the learning tasks was counterbalanced across participants. The total number of trials can be summarized as:

   30 participants x

   3 learning tasks x

   3 trials

   = 270 trials.

   Between each trial, participants were allowed to take a short rest if needed. After 3 trials, we instructed participants that they could continue playing for another 10 minutes if they wanted to. During the study, experimenters only interfere when a technical problem occurred.

4. Assessment (5 minutes): Participants answered two subjective assessment questionnaires regarding their experiences with the gamified/non-gamified system. The first questionnaire was comprised of 14 statements, asking participants to indicate their level of agreement on a 7-point Likert scale (7 being Strongly Agree). We adapted these items from the Intrinsic Motivation Inventory [73], GEQ [54], and the attitudinal survey used in [29]. The second survey intended to assess the general attitudes toward Points, Badges, and Leaderboard in the experimental group. The second survey consisted of nine items on a 7-point Likert scale with 7 as Strongly Agree.

After both sessions ended, we performed a follow-up semi-structured interview, to collect subjective feedback.

Figure 7.3 shows the engagment model of this study. Yellow-highlighted objects show the updated part of Study 5 from Study 3.

7.2 Results and analysis

We analyzed the data using repeated measures ANOVA and post-hoc comparison with Bonferroni correction to study the effects of the game elements. Our analysis reveals several key

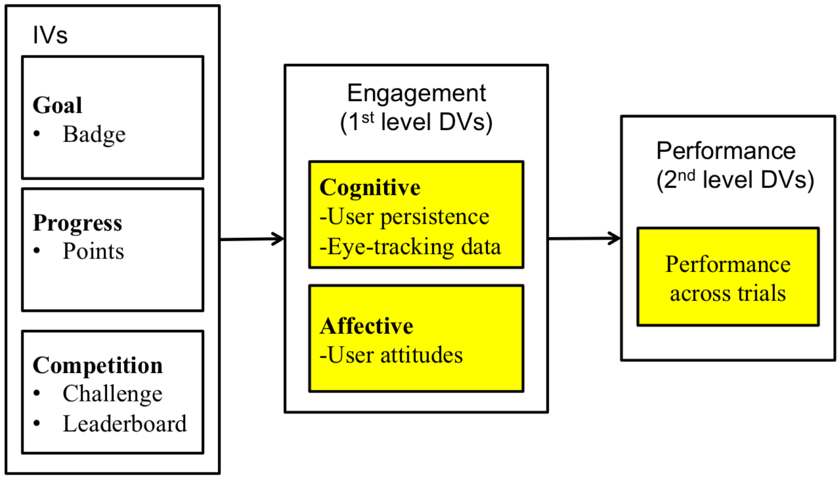


Figure 7.3: Study 5's engagement model

findings, described as follows.

7.2.1 Points, Badges, Leaderboard Increased Persistence

We hypothesized that participants using the experimental version tend to commit more time and efforts to the tasks. To evaluate this hypothesis, we counted the number of users who performed extra rounds of exercises in each task, aside from the three mandatory trials. Since the standard learning tasks that we provided are not designed for fun, we expected users to quit after meeting the minimum requirement. And thus the number of users who performed extra trials becomes an indication of an increased desire to focus consistently on the learning tasks.

ANOVA results showed that experimental group has a significant effect on persistence, with 20 users (22%) in the experimental condition performing extra trials, compared to 7 (8%) in the control condition ($F_{1,178}$=6.55, p<0.01) (See Figure 7.4). For those who did conduct more exercises, the average number of additional trials completed was 2.88 (SD=0.48) in the experimental condition, compared with 1.89 in the control condition (SD=0.51), although the effect was not significant.

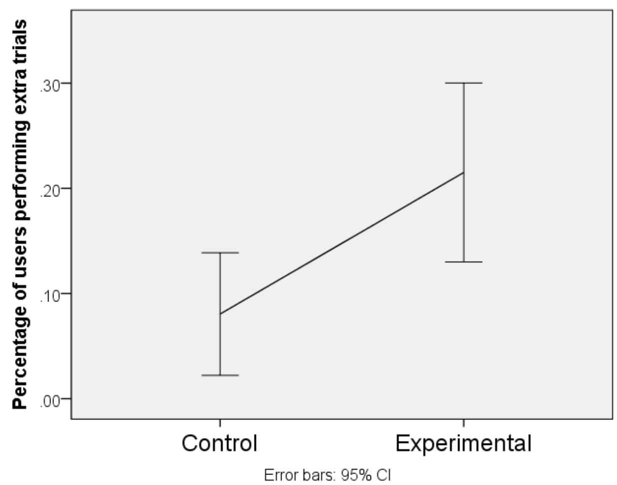


Figure 7.4: Comparing persistence between control and experimental group. Data showed significant differences (p<0.01).

We further analyzed whether the nature of tasks (memory vs. problem-solving) influenced persistence. ANOVA results suggested that the effect of game elements on persistence in memory tasks was not significant ($F_{1,87}$=2.60, p=0.116), with five users in the experimental condition and one in the control condition putting in extra efforts. By contrast, the difference in problem-solving tasks was significant ($F_{1,89}$=4.39, p<0.01), with 15 users in the experimental group and six users in the control condition. The insignificance in memory tasks may be due to the nature of the task which requires a large cognitive load and thus users are less likely to persist when compared to problem-solving tasks. Indeed, some participants in the experimental group reported in the final interview that they felt "exhausted" memorizing items and could not continue after three trials, even though they would like to.

These results suggested that game elements can increase the likelihood that users will devote more efforts to learning, though the degree of impact varies with the nature of the learning task.

7.2.2 Points, Badges, Leaderboard Improved User's Attitudes

We hypothesized that users of the experimental group would generally react more positively towards the learning tool and task experience than those in the control group. To evaluate this hypothesis, we collected participants' responses in the Q1 attitude questionnaire. We aggregated each person's level of agreement on 14 statements (on a 7-point scale) into an average score, given a high level of internal reliability (Conbrach α of 0.86).

ANOVA comparison showed that the game elements had a significant effect on user attitudes and perceptions, with users in the experimental condition enjoying the learning experiences more (M=5.01, SD=0.17) than those in the control condition (M=4.45, SD=0.21) ($F_{1,56}$=4.46, p<0.05) (see Table 7.2).

We further looked into whether the nature of the task (memory vs. problem-solving) made a difference. We found that users in the experimental group scored significantly higher than the control group, in both memory ($F_{1,26}$=7.64, p<0.01 - 4.71 vs. 3.72) and problem-solving ($F_{1,27}$=6.29, p<0.01 - 5.50 vs. 4.86) sessions. However, when considering only the attitude of the experimental groups, problem solving tasks received a significantly higher score (M=5.5) than the memory tasks (M=4.71) ($F_{1,27}$=9.88, p<0.01). Our interview revealed that many players perceived memory tasks to be more cognitively demanding and difficult, and thus were relatively less enthusiastic about them. Overall, PBL can improve one's attitudes in learning even though the tasks are perceived to be difficult or unpleasing, but the initial motivation and perception of the learning tasks did have an influence on its effect.

7.2.3 The Leaderboard Ranked Highest in Preference

We were also interested to determine which of the three game mechanics (Points, Badges, Leaderboard) has the biggest impact from the users' viewpoint. We gathered participants' preferences, in terms of interest, attention, and motivation of use, regarding Points, Badges, and Leaderboard respectively via the questionnaire Q2. The 7-point scale scores for each of the game mechanics were aggregated into an average score, given the high levels of internal

Item	Control	Exp.
#1 I enjoyed the activities.	5	5.43
#2 I think the activities are fun.	4.47	5.11
#3 I think the activities are easy to understand and perform.	5.67	6.04
#4 I was concentrated during the activities.	4.6	5.57
#5 I did not take the activities seriously (R).	5.1	6
#6 I paid little attention to my performance (R).	3.93	5.18
#7 I want to play these activities again.	3.07	3.89
#8 I felt bored (R).	4.67	5.18
#9 I felt the tasks were too difficult (R).	4.8	4.68
#10 I felt challenged.	4.13	4.64
#11 I put in effort to perform the activities.	4.97	5.64
#12 I think I am pretty good at the activities.	3.8	4.36
#13 I am satisfied with my performance.	4.03	4.54
#14 I felt my performance has improved after a few trials.	4.27	4.46
Total	**4.45**	**5.01**

Table 7.2: Questions and results of the subjective questionnaire in a seven-point Likert scale (7-Strongly Agree). Results showed that the experimental condition encourages significantly better attitudes ($p<0.05$). Note: R=reverse items; scores have already been reversed.

reliability (Conbrach α ranging from 0.94 to 0.97).

ANOVA showed that there was a significant difference ($F_{2,87}=10.47$, $p<0.001$) in the perception of the three elements, with Leaderboard scoring 5.74, Points scoring 5.35, and Badges scoring 4.09. A posthoc test with Bonferroni correction confirmed the significant differences between Leaderboard and Badges ($p<0.001$), and between Points and Badges ($p<0.001$), but not between Leaderboard and Points ($p=0.887$) (See Figure 7.5). Similar results were observed when split into memory and problem-solving tasks. Our eye-tracking data were coherent with the questionnaire results, with the distribution of fixation being: Points (35.63%), Badges (8.73%) and Leaderboard (55.64%).

The fact that Leaderboard and Points were rated relatively high was anticipated, as both elements may trigger explicit or implicit competition. What surprised us was the relatively low score of Badges. We expected that the goal setting capability of Badges would more significantly improve user experiences. However, our results contradicted prior work which stated that Badges have significant motivational effect on users (e.g., [31]). Post-study interviews ruled out usability or aesthetical problems with our Badges. One possible explanation is that the effect of the Leaderboard and Points hindered the goal setting of Badges, i.e., the sense of achievement of getting better than others is stronger than the sense of achievement when earning new Badges. For example, a participant commented - "*Being number one in the Leaderboard is my top goal...maybe that's why I feel less excited to compare my badges which I give lower priority.*". In other words, in our tasks which are more utilitarian-oriented, they encouraged a substantially more cognitive involvement (getting better performance) than affective involvement (collecting badges). Thus the affective components of the system (i.e., Badges) might be chosen to be ignored by users in favor of cognitive involvement. In other words, cognitive involvement may overrule the affective information when people exercise tasks with a determined goal.

Another related explanation is that our tasks were mainly point-based, and thus the goal was relatively clear from the beginning (to get the highest points). As a result, users may not need Badges to help clarify the goals.

This implied that Badges may be more effective in the case when the goal is less obvious or the performance is harder to measure, and when the assigned task is more hedonically-oriented, i.e. encouraging affective involvement. This also suggested that Badges may be very useful in exploratory types of learning applications in which exploration and discovery (information seeking) are the primary learning methods. For example, Badges may be useful in an application for learning chemistry where learners are asked to discover different types of chemical reactions, since Badges can guide users to relevant learning objectives which may not be obvious. Providing Badges for random discovery can also provide a sense of surprise. Last but not least, Badges may be suitable for encouraging the exercise of certain behaviors. In comparison, Points and Leaderboard are effective particularly when rating and ranking are meaningful to users. For example, if participants see the points and rankings as an indicator of their IQ level, they might be more motivated to climb up the Leaderboard.

7.2.4 Points, Badges, Leaderboard Improved Performance

We hypothesized that the experimental group would achieve better task performance than the control group. Specifically, we expected that the experimental group would achieve greater

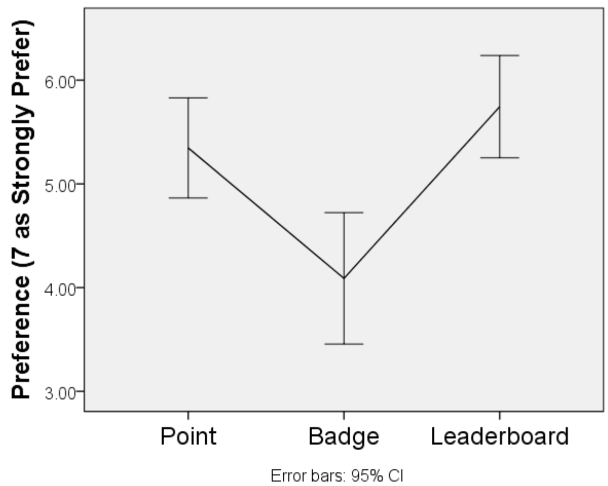


Figure 7.5: Comparing preferences between points, badges, and leaderboard. ANOVA tests showed significant differences ($p<0.001$). Posthoc with Bonferroni correction confirmed differences between the Leaderboard and Badges ($p<0.001$), and between Points and Badges ($p<0.001$), but not between the Leaderboard and Points ($p=0.22$)

progress via repeated practice. Task performance data showed that both experimental and control groups did improve their ability over the course of multiple trials. ANOVA revealed statistical differences in the performance of problem-solving tasks (see Table 7.3 and Table 7.4): control-TOL ($F_{2,40}=4.99$, $p<0.01$), control-WS ($F_{2,42}=3.69$, $p<0.05$), experimental-WCST ($F_{2,39}=3.93$, $p<0.05$), experimental-TOL ($F_{2,42}=6.51$, $p<0.001$), and experimental-WS ($F_{2,42}=15.98$, $p<0.001$); but not in memory tasks. Although the control group significantly improved their results between the first and third mandatory trials in TOL and WS, but not as much as in the experimental condition - WS achieved a 60% improvement in the experimental group and a 33% improvement in the control group; TOL achieved a 22% improvement in the experimental group and a 16% in the control group. In addition, only the experimental group achieved significant improvement in WCST.

We also asked participants to rate their perceived improvements for each task on a scale of seven. Most participants in the experimental group felt that their problem-solving skills

		Control			
Task	Measure	1st trial	2nd trial	3rd trial	**Sig.**
YNR	Correct responses	18.01	18.85	21	-
CR	Correct responses	3.23	3.54	3.62	-
FR	Correct responses	7.31	7.92	7.85	-
WCST	Correct matches	45.64	48.07	46.86	-
TOL	Time[1] ; # of moves[2]	151.47; 12.76	123.49; 12.69	126.59; 12.94	**[1]; -
WS	# of failed attempts	3.60	3.20	2.44	*

Table 7.3: Users' mean performance on the three trials in the control group. ANOVA tests showed significant differences between trials in the control group.

		Experimental			
Task	Measure	1st trial	2nd trial	3rd trial	**Sig.**
YNR	Correct responses	18.81	19.31	18.56	-
CR	Correct responses	3.56	3.5	3.75	-
FR	Correct responses	8.44	9.06	8.81	-
WCST	Correct matches	43.21	46.43	47.72	*
TOL	Time[1] ; # of moves[2]	152.87, 13.41	118.24, 12.48	119.525, 13.16	***[1];-
WS	# of failed attempts	3.67	2.34	1.47	***

Table 7.4: Users' mean performance on the three trials in the experiment group. ANOVA tests showed significant differences between trials.

got better over time (M=5.12), but they were not as confident about their memory skills (M=4.03). By contrast, participants in the control group did not think that their problem-solving skills (4.56) or memory skills (3.89) had been any different.

We further analyzed users' performance over the extra trials. For those who did performed additional trials in the memory tasks, no one in either the experimental or the control group made any progress. On the contrary, all users who did extra work in the problem-solving tasks achieved better results compared to the 3rd trial, whichever group they were in. This suggests that persistence and practices did improve people's problem-solving skills.

In the pretest-posttest comparison, we did not find any significant differences between the control and the experimental group. This indicates that although the task performance had significantly improved, the transfer of knowledge may require more practice.

In summary, these results confirmed that the embedded game elements improve performance through higher concentration on the task at hand and more persistent commitment to learning. They also encourage people to engage in the tasks more actively.

7.2.5 Game Engagement and Achievement Motivation Theory

We identified different types of learners via coding the interview transcripts, which were consistent with the Achievement Motivation Theory [4] that makes a distinction between two personality traits: Need to Achieve (Nach) and Need to Avoid Failure (Naf). Participants with high Nach (High-Achievers) tend to be better motivated by the existence of PBL, since PBL generates a perception of progress and competition with self and others. They are less afraid of failure and thus are more likely to take on more challenges. They were coded with common schemes such as "*I tried my best*", "*It's challenging*", "*I do not want to lose to others*", "*The points and leaderboard keep me informed of where I am*".

Participants with high Naf (Low-Achievers), on the other hand, appear to be easily discouraged. For example, once low-achievers consider catching up with the top players on the Leaderboard to be impossible, they may quickly lose interest in the task. Low-achievers may form such a judgment before they even start the task (they looked at the top score in the Leaderboard and felt it unachievable), or as soon as they completed a certain task (they felt that they could not do any better and were satisfied with their current points). In other words, low-achievers' fear of failure outweighs their desire of success. They were commonly coded with schemes such as "*The top score is too high*", "*this game is too difficult*", "*I stopped looking at the Leaderboard*".

We can further categorize the high-achievers into two types: Mastery-oriented and Ego-oriented [4]. Mastery-oriented high-achievers were users who pursue self-improvement. Regardless of the presence of PBL, these users constantly motivate themselves to forge ahead. Ego-oriented users seemed to benefit more from PBL which provides extrinsic stimulation by enabling them to compare their performance with others'. These users were more interested in looking at the Leaderboard to ensure they were among the top players.

According to the Achievement Motivation Theory, the reason why the participants in our experiment showed clear distinction in their personality is due to two situational factors: the probability of success and the importance of success. Probability of success refers to the likelihood of achieving success. In our case, the ranking on the leaderboard gave low-achievers the impression of an unreachable goal, while high-achievers viewed it as an opportunity to conquer. The importance of success refers to the intrinsic meaning of success to users. In our study, since the fundamental learning tasks were somewhat like an IQ test, participants may interpret their task performance as a reflection of their intelligence (IQ) level. As a consequence, high achievers can be greatly motivated, whereas low-achievers may be scared by the learning objectives. Low-achievers were afraid to see that their scores were around the average or lower levels, and this is apt to impair their self-confidence and damage their social perception of others [4].

This result suggested that PBL should be carefully designed to accommodate low-achievers. For example, by supporting a local ranking feature (ranking around users with the same level of skills), low-achievers might be less intimidated. It is also worth investigating whether providing alternative self-improvement indicators such as Badges might work better for low-achievers.

7.3 Discussion

In summary, our study confirmed that PBL can significantly facilitate learning. Specifically, the use of PBL improved persistence and attitudes in abstract tasks. The use of gamification does not directly teach people anything; instead, people's improvement come from longer engagement with and higher concentration on the tasks. PBL gamification may also encourage people, whether it is to compete with self or others, or to seek better solutions themselves instead of merely repeating the same method. However, the mechanics need to be carefully designed. The learning tasks being too hard or too easy may both reduce their impact. Last but not least, the design of PBL should take special consideration when targeting low-achiever users.

Our study also demonstrated how abstract tasks adapted from cognitive psychology can be used as experimental tools for learning. Using standard abstract tasks can simplify experimental design by eliminating possible confounding factors such as the choice of the subject or the contents of the learning materials. In any case, it was our intention to show the merits of abstract tasks as a complementary method, rather than a replacement of existing "in-the-wild" testing methods. Such abstract tasks can be employed first for preliminary, controlled evaluation before a large-scale field test. Both will be needed to provide a solid replication of results.

Some limitations may affect the results. First, this was a short-term study. As a result, we were not able to observe any improvement on knowledge transfer in the pretest-posttest comparison which might need a longer time to achieve. However, as a result of this limitation, we used other metrics to complement our results such as 1) participant's progress through the different trials of the game (as they perform a task multiple times); 2) people's willingness to perform more (extra) trials, especially in the gamified mode; and 3) interview data to find out explicitly if participants perceived that they have improved. This data also provides alternative evidence of user improvement.

Second, there might be possible carry-over effects for some participants, since they can be in the control-group for one set of tasks, and in the experimental group for another based on their pretest performance. Foreknowledge of the existence of PBL may affect the validity of the results (not strictly a between-subject design). To assess the impact of the carry-over effect, we excluded the data of participants who were in both the experimental and control group and re-ran the analysis. We received similar results; for example, we found significant differences in attitudes between the experimental and the control groups ($F_{1,25}$=6.42, $p<0.01$) and a significance in persistence between the two conditions ($F_{1,105}$=5.02, $p<0.05$). This ensures that the carry-over effect did not have any significant impact on the validity of this paper's results.

Third, the improvement in performance in TOL may need to be treated cautiously, since only the 2nd trial had significant improvement over the 1st trial, with the 3rd trial slightly worse than the 2nd in both groups. One possible explanation is the learning effect - being familiar with the interface and the task. However, we verified that this was not the case by checking the performance of users who conducted additional trials. Our analysis found that these users made significant progress through extra practices, i.e., 8.7% improvement in control group and 19.5% improvement in experimental group on average compared to their results in the 3rd trial. This suggests that users did achieve progress over the course of multiple trials. The slight drop in performance in the 3rd trial might be due to coincidental factors, such as fatigue or user's contentment in their improved performance on the 2nd trial.

Chapter 8

STUDY 6: EMPIRICAL STUDY IN FULL-BODY GAMES (I)

We have evaluated the effect of game engagement in gamification. We now shifted our focus to full-body games, given its potential for health and rehabilitation purposes. Full-body game engagement is also associated with the concept of embodiment. We believed there are still rooms for fundamental studies to improve full-body game engagement.

Study 6 focused on designing engaging gestures, concerning the interaction in the framework (see Figure 8.1).

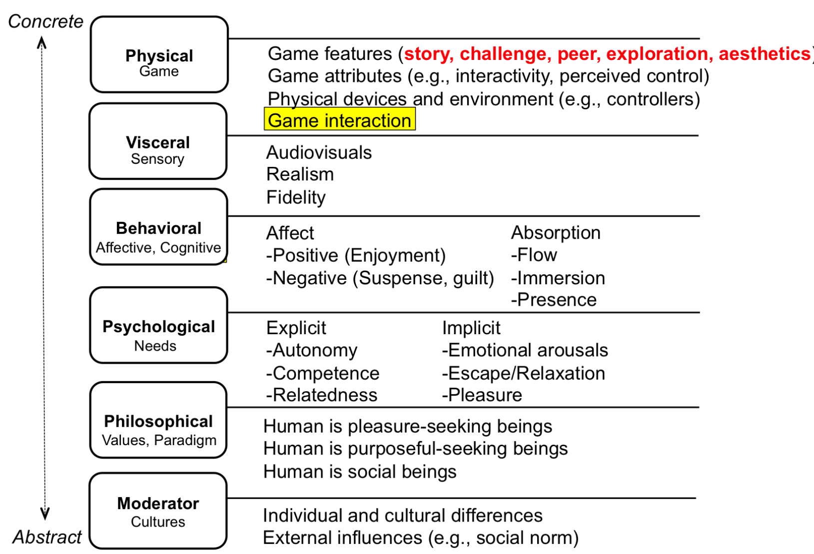


Figure 8.1: Study 6: Evaluating full-body game gestural interaction

Full-body based interaction (e.g., Kinect) has enabled more natural and intuitive input for video games. However, game gestures developed by designers may not always be the most suitable gestures for players. Indeed, players have reported difficulties in playing some full-body based games, particularly in interaction-intensive games (e.g., First Person Shooters/Action/Adventure) where several actions/commands may have to be executed at or nearly at the same time [e.g. 41]. Thus one key challenge in designing engaging game gestural interfaces lies in defining suitable, efficient gestures that enable players to effectively perform multiple game actions/commands simultaneously and with ease.

Several studies in relation to full-body game interaction have been conducted [e.g., 50, 84], but few studies have considered the intense-dynamic nature of game environments in general. When a player's hand is occupied with "Shooting Zombies", which other body parts and gestures might the player prefer to perform simultaneous actions such as "Reload" or "Use First Aid Kit" with. Since a literal "Jump" or "Climb" action can be tiring, is it likely that users will prefer a less tiring, more efficient gesture? What gestures would veteran gamers and non-gamers devise or envisage to enhance their interaction experiences?

To investigate these potentials, three user studies were conducted. In the first study, to explore general user preferences of game gestures, we used a user-elicitation approach asking participants to define their preferred gestures for different game actions/commands. We found a high consensus (agreement score) between participants' gestures as most participants defined physical gesture (mimicking real-world actions) with 1-hand as the most preferable input body modality. We also found a difference in preferences between gamers and non-gamers.

In the second study, to also consider simultaneous use of gestures where physical gestures may not always be possible, we asked participants to rate the suitability of different body parts (one and two hands, one and two legs, head, eyes, torso) for each game action/command. This second study was intended to help designers consider a set of suitable and alternative body parts, since an alternative body part may be needed to execute other simultaneous gestures while a certain body part is already occupied. Through the study, we identified a set of suitable and alternative body parts and gestures for different game actions/commands.

In the third study, to develop a simultaneous gesture set, we initially asked three participants to define their preferred gesture set using the user-elicitation approach. However, we found that there was little consensus among participants. In addition, participants mentioned that it is difficult to imagine possible combinations of gestures. To assist participants, we adapted the original user-elicitation approach and introduced a novel choice-based elicitation approach. We found that this approach has a positive effect in assisting participants to discover and create suitable gestures, which resulted in a consensus set of simultaneous game gestures. Based on the three studies' findings, we highlight potentially useful design guidelines.

8.1 *Methodology (I)*

8.1.1 Design

To first explore player's preferences in game gestures, a user-elicitation study was conducted, i.e., participants were asked to define single gestures for each common game action/command. Agreement scores, frequency of use of body parts, gesture types, and subjective assessment were analyzed.

8.1.2 Participants

Twelve university students (all males, M=22.1 years) were recruited. Regarding game experience, five of the participants regularly played games in both PC (mouse+keyboard) and consoles (game controller) with more than 15 game hours per week (veteran gamers). Regarding experience with Kinect or other motion game gestures, only seven of the participants (including all five veteran gamers) reported having some prior experience (one to three times). All participants were right-handed and each was paid $10.

8.1.3 Events[1]

We derived events from various genres of Kinect games including *Gunstringer* (First Person Shooters), *Blackwater* (First Person Shooters), *Forza 4* (Racing), *Kinect Joy Ride* (Racing), *Kinect Adventure* (Adventure), *Kinect Rush A Disney Pixar Adventure* (Adventure and Role-playing), *Rise of Nightmares* (Adventure and Role-playing), *Kinect Sports* (Fighting and Sport), *Virtual Tennis 4* (Sport), *London 2012 Olympics* (Sport). This list of games is in no way exhaustive but these games do cover various genres of motion gaming, thus implying that these games can serve as a good representative starting point. We also included a few actions not included in current Kinect games but we considered them to be typical in video games such as "Stealth Walking" and "Steal". We left out a few repetitive actions such as "Use First Aid Kit" and "Use Power Up" and generalized them into one common action, e.g., "Use Item". There are some actions such as "Catch Ball" or "Jump" where the resulting gestures could be obvious, but including them in our study enables us to observe any specific behaviors (e.g., 1-leg vs. 2-leg).

A total of thirty-two actions and commands were derived. The complete set of events includes first person shooter's actions (Shoot, Reload Gun, Viewpoint Change), racing actions (Drive, Accelerate, Shift gear), adventure/role-playing actions (Climb, Walk, Stop Walking, Run, Jump, Stealth Walking, Steal, Slash, Open Chest, Open Door, Pick Item, Use Item, Push Box, Shake Item), sport/fighting actions (Hitting Baseball, Catch Ball, Throw Ball, Row Boat, Roll Bowling, Racket Hitting, Kick, Guard), and system commands (Zoom-In, Zoom-Out, Open Menu, Select Menu Item).

8.1.4 Procedure

Our study design used a user-elicitation approach similar to that of Wobbrock et al. [123] and Ruiz et al. [100]. At the start of the experiment, participants were asked to define game gestures for 32 game events in randomized order. To identify the most preferred gesture and

[1] For simplicity, the term *"events"* will be used in place of *"actions and commands"* throughout this paper.

reduce any ordering effects, participants were allowed to choose the same gesture for multiple events. Each event was displayed along with the name of the event and a target scene on a large display. Target scenes were created by using Visual Studio 3D Toolkit to represent an effect (e.g., a treasure chest being opened, an opponent being slashed), and participants were asked to perform gestures to cause (trigger) the effect. Target scenes were not taken from animated screenshots of any existing video games, as we tried to keep the target scenes independent of any particular game, which might otherwise influence the resulting gesture. Some target scenes were not animated, e.g., “Stealth walking” or “Drive”, as these scenes require the participants to see the actors in a third-person perspective to show the effect clearly, which may influence the participant’s defined gestures; so we simply communicated the effect using instructions along with static images containing the interaction medium/context of the events (see Figure 8.2).

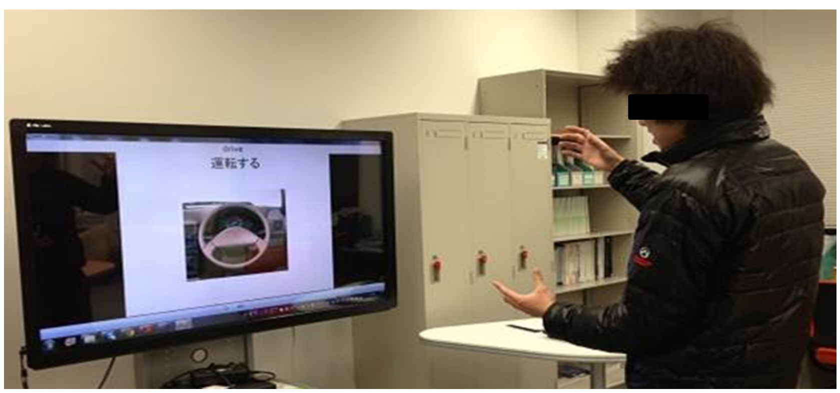


Figure 8.2: In Experiment I, participants were asked to perform gesture for different common game events.

During gesture definition, participants were instructed to think-aloud while performing their gestures, confirm the start and end of their performed gesture and describe the corresponding rationale. Participants stood approximately 1.8 meters away from the display while performing gestures. The experiment was audio and video recorded for later analysis. As with common elicitation studies, we did not want participants to take recognizer issues into consideration, i.e., to remove the *gulf of execution* [53] between the user and the device, and to observe the users' best possible gesture without users being affected by the recognition ability of the system - similar to the rationale described in Wobbrock et al. [123] and Ruiz et al. [100].

A similar evaluation method to that of Wobbrock et al. [123] was used: "The gesture I performed is a good match for its purpose"; "The gesture I performed is easy to perform"; "The gesture I performed is tiring". The questionnaire followed a 7-point Likert scale style with 7 as "strongly agree". Participants evaluated the gestures immediately after each performed gesture to assist recall efforts. To improve consistency of the evaluation, after all gestures were performed and evaluated, participants were allowed to double-check their evaluation scores for each gesture, and when needed, to look at the video or readjust their scores if needed. In our study, only two participants revised their scores but without any dramatic changes (i.e., both participants changed their scores for only two gestures by the scale of one, e.g., 6 to 7) and did not affect our results. The experimental session took around 1-hour.

8.2 Results and analysis (I)

Agreement scores, frequency of use of body parts, gesture types, and subjective assessment were analyzed. A total of 384 gestures were collected.

8.2.1 Agreement score

Wobbrock et al. [123]'s method was used to investigate the degree of consensus for each game event. The calculation of agreement score is as follows:

$$A_r = \sum_{Pi} \left(\left| \frac{P_i}{P_r} \right| \right)^2$$

Pr represents all gestures performed for event r and Pi is a subset of identical gestures from Pr. Ar ranges 0 to 1. Gestures were considered identical if they have similar trajectories and poses. For example, for the agreement score of the event "Walk", a total of 12 participants and their corresponding 12 gestures were considered. This event was divided into 4 groups of identical gestures. The size of each group was 7, 3, 1, and 1. Therefore, the agreement score for "Walk" is:

$$A_{walk} = \left(\frac{7}{12}\right)^2 + \left(\frac{3}{12}\right)^2 + \left(\frac{1}{12}\right)^2 + \left(\frac{1}{12}\right)^2 = 0.42$$

Figure 8.3 shows the agreement score for each game event. The overall agreement score was 0.37, which is slightly higher than Wobbrock et al. [123](0.32 and 0.28). Regardless of the high overall agreement score, system commands including "Open Menu", "Zoom-In", "Zoom-Out" achieved relatively low overall agreement scores (0.126).

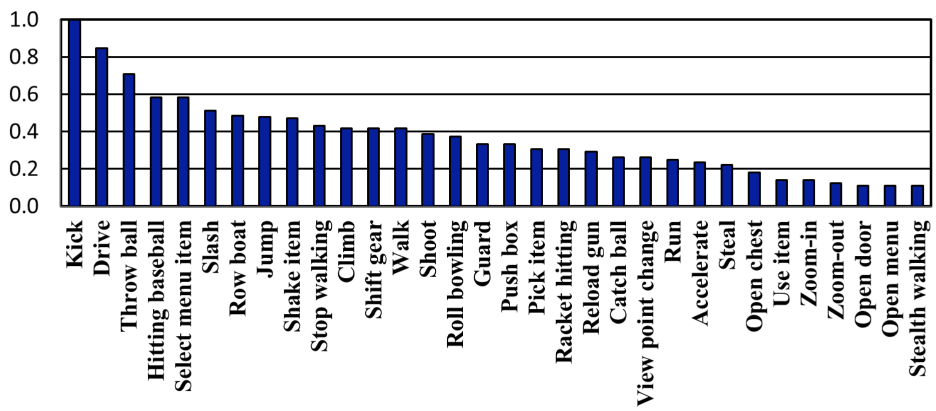
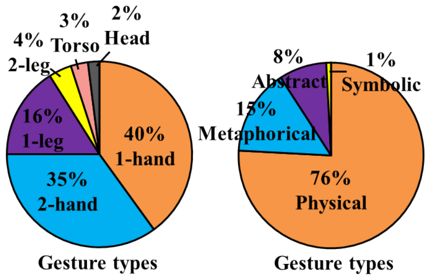


Figure 8.3: Experiment I - Overall agreement scores for events sorted in descending order.

8.2.2 Use of body parts and gesture types

Figure 8.4 shows the use of body parts. It is not surprising that one-handed gestures were the most preferred (40%), followed by two handed (35%), one leg (16%), two legs (4%), torso movement (3%), and head (2%). However, this pattern was not the same across all actions. For navigational events such as "Run", "Walk', "Stop Walking", most participants preferred leg gestures. For the "Viewpoint Change" event, participants preferred head input



	Physical	Metaphor	Abstract	Symbolic
Normal / non-gamers	66%	32%	40%	0%
Veteran gamers	34%	67%	60%	100%

Figure 8.4: The proportion of body parts used and gesture types (left) and veteran gamers vs. non/normal gamers (right).

with a few using the torso. For the commands "Select Menu Item" and "Open Menu", many participants preferred hands, however head input was used by some.

Regarding gesture types, it is important to highlight the difference in gesture preferences between veteran gamers and normal/non-gamers. We classified the defined gestures into four broad types [123]: physical (direct manipulation), symbolic (making an "OK" gesture), metaphorical (using a foot to "double click" items), and abstract (arbitrary mapping). The classification was conducted by the authors. To improve reliability, an independent rater performed the same classification with high interrater reliability (Kappa=0.932, $p<0.001$).

As shown in Figure 8.4, most defined gestures resembled real-world gestures (physical gestures). On the other hand, non-physical gestures were mostly performed by veteran gamers, e.g., for the action "Jump", non-gamers tend to prefer using an actual "Jump" action while veteran gamers preferred a less tiring gesture (e.g., raising a leg slightly above the ground). In any case, the found differences between veteran and novice gamers should be treated suggestively due to the small sample size of veteran gamers (n=5).

8.2.3 Subjective assessment

Participants evaluated their defined gestures based on the Likert 7-scale rating (7 as strongly agree). The mean value is "The gesture I performed is a good match for its purpose?" (5.84); "The gesture I performed is easy to perform" (5.90); and "The gesture I performed is tiring"

(5.51). The correlation between participants' good match ratings and agreement scores was found to be significant (r=0.746, p<0.01), suggesting that a good match rating is a good indicator of the consensus between users of the gestures performed.

8.3 Discussion (I)

Individual difference. Our study achieved a relatively high agreement score compared to past works. The high agreement score was due to participants often mimicking their real-world gestures (e.g., kick) for performing game events. However, there were also some actions that exhibit slight variations which reduce its agreement score such as "Racket Hitting" - four different gestures were produced - two-arm swing, one-arm swing, wrist-flicking, or palm-gesture for "Racket Hitting". The variations in gestures reflect individual differences - some participants prioritize efforts required while some attempts to maximize the correspondence between game events and real-world gestures. One individual characteristic we found to affect gesture preference is gaming expertise. More specifically, gamers tend to define a more symbolic, efficient unique gesture library, while less-experienced gamers tend to define a more straightforward, physical gesture set. Possible explanations include the observation that veteran gamers tend to play games for a longer period of time, thus they prefer more efficient, less tiresome gestures to prolong their duration of play. Another possible explanation is that veteran gamers may tend to be more motivated to overcome challenges, thus they prefer more efficient gestures to competitively engage in those challenges [83, 88]. Last, it may simply be due to the overall greater game experiences of veteran gamers, thus leading them to define more efficient game gestures.

To accommodate this difference, the idea of accommodating multiple gestures per game event has support [84]. For example, it might be useful for designers to incorporate two possible gestures for one event - a physical gesture + a more efficient, symbolic gesture. Such accommodation (e.g., shortcut commands) has already been implemented in a desktop-based system facilitating novice and expert users' needs and thus implying that it could be useful for full-body interaction. Aside from gaming expertise, designers might also need to consider varying levels of player motivation [77] and also consider that the same gesture may not be valid for all users, places [40, 79] or situations (e.g., gamepad might be preferred over a long session of gameplay).

2D vs. 3D interaction. Regarding system commands such as "Zoom-in/out" and "Open menu", these commands achieved relatively low agreement scores. Although gestures for these commands were well-defined in 2D interfaces, they do not appear to be transferable to 3D interfaces. This indicates that designers might need to consider developing new 3D gestures for similar commands. Perhaps due to the larger degree of freedom when compared to 2D interaction paradigms (e.g., mobile interactions), the elicitation resulted with diverse gesture definitions. It may be beneficial for designers to investigate the right degree of freedom for users on similar system commands.

Limitations. Experiment I has several limitations - first, alternative body parts for each game event were not adequately explored in cases where the preferred body part was occupied with the task at hand. For example, it may not be possible to use a hand for both "Drive" and "Shift Gear". Alternative body parts should be identified to assist designers during the assignment of gestures; second, few possible simultaneous game gestures were discovered, i.e., participants mostly defined gestures based on their real-world actions but such actions may not always be applicable in highly interactive game situations and gestures more applicable to digital gaming might thus be overlooked. These issues were addressed in Experiment II (alternative body parts) and III (simultaneous gestures) respectively.

8.4 Methodology (II)

8.4.1 Design

In Experiment II, we seek to identify suitable and alternative body parts. We asked participants to rate the suitability of each of the body parts for each game event.

Figure 8.5: In Experiment II, we asked participants to perform gestures using different body parts. In the figure above, we asked participants to “Shoot” using head (the participant moved head forward to “Shoot”) and other body parts as well.

8.4.2 Participants

Another twelve university students (1 female, M=22.4 years) were recruited. Regarding game experience, six were veteran gamers who regularly played games on both PC and consoles with more than 15 hours per week. Regarding experience with Kinect or other motion game gestures, only eight of the participants (including all veteran gamers) had prior experience with full-body games (two to three times). All were right-handed. They were paid $10. The only female participant had no experience in video games. In our study, we did not observe any gender-specific differences between the female participant and other male non-gamer participants.

8.4.3 Events

For better coverage, 8 more game events were added for analysis including adventure/role-playing actions (Crawl, Roll Body), sport/fighting actions (Dodge, Headbutt, Punch), racing actions (Break, Boost), first person shooters actions (Hide).

8.4.4 Procedure

Participants were asked to define possible game gestures for 40 game events using different body parts in randomized order. These body parts were derived from our pre-study where we asked eight participants (6 males, M=21.3 years) to nominate the possible body parts for performing mid-air gestures. We grouped similar body parts (e.g., wrist/arm/finger → hand) and identified seven main body parts (one and two hands, one and two legs, head, eyes, torso) to be used in this study. The group classification was achieved with high interrater reliability (Kappa=0.985, $p<0.001$).

As opposed to performing *only one* gesture using any body part for each game event (Experiment I), Experiment II asked participants to perform *one gesture per one body part* (thus seven gestures per game event) and rate the suitability of each body part - "The body part I used is suitable". The questionnaire followed a 7-point Likert scale style with 7 as "strongly suitable".

Participants were instructed to skip when they felt that no suitable gestures could be performed using a particular body part. For each body part, participants were allowed to use different subparts (e.g., wrist/thumbs of the hand, elbow of the arm, feet of the leg). Similarly, we instructed participants that eye gestures could be any interaction performed using eye movements or fixations. Participants were also instructed that they were allowed to touch another body part (e.g., touching the left shoulder with right hand) where the gesture would be counted as a hand movement if the hand is the initiator. The setting was similar to that used in Experiment 1. The experimental session took around 1.5-hours.

8.5 Results and analysis (II)

The agreement score, suitable and alternative body parts, and gestures were analyzed and reported.

Agreement score

We used a similar approach to Experiment I to calculate the agreement score. Figure 8.6 shows the agreement score for each game event. Gestures were considered identical if they have similar trajectories and poses. The mean agreement score was 0.19, which was much lower than Experiment I (0.37). This was expected due to the design of Experiment II which asked participants to perform each gesture using different body parts.

Many events (e.g., Kick, Drive, Shoot) that achieved high agreement scores in Experiment I have lower agreement score in Experiment II. For example, there were a total of 11 different gestures performed by participants to “Shoot”, e.g., using one hand/two hand as a gun, using head movement, using the whole arm as a gun, using a kick-gesture.

On the other hand, events such as “Headbutt”, “Racket Hitting”, and “Catch Ball” achieved high agreement scores, given that most participants agreed that the most natural way to perform such actions is to imitate real-world actions.

Comparing Experiments I and II, the largest group of defined gestures was consistent, i.e., 31 out of 32 events in both Experiment I and II have the same largest group of defined gestures. For example, in both Experiment I and II, the most preferred gesture for “Shoot” is a one hand gun-gesture.

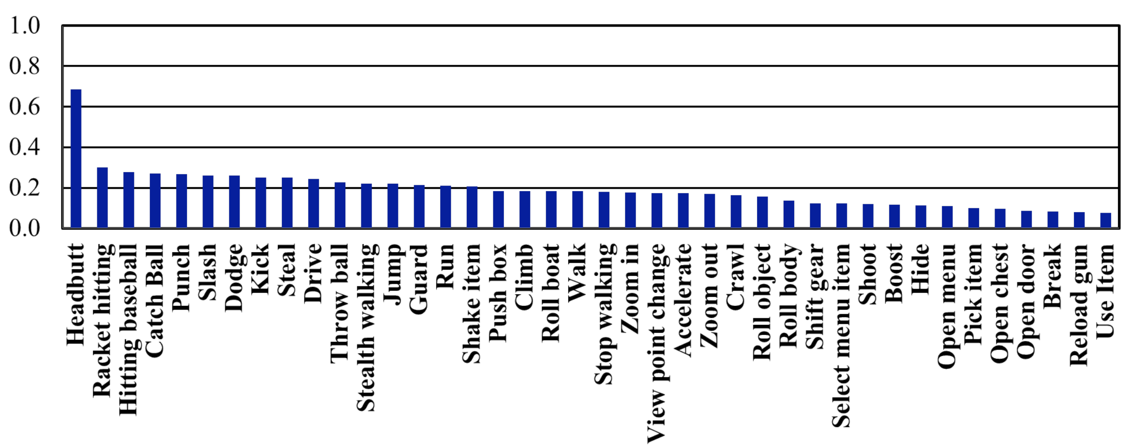


Figure 8.6: Experiment II - Overall agreement scores for events sorted in descending order.

Suitable and alternative body parts and gestures

We analyzed the most suitable body parts and alternatives based on the participants' subjective assessments of "The body part I used is suitable". The most suitable body part for each event was selected based on the highest mean score. We also analyzed alternative body parts in case the most suitable body part was already occupied with the task at hand. Alternative body parts were selected on the grounds that they did not conflict with the most suitable body part and they presented with the highest mean score among all the alternatives. In cases where two body parts achieved equivalent highest mean scores and posed no conflict in body part usage, both were highlighted as the most suitable body part and no alternative was highlighted (e.g., Roll body).

Figure 8.7 shows the suitable and alternative body parts for each game event. To apply Figure 8.7, one can consider the score to determine the suitable combinations. For example, if designers were to design combined gestures of "Drive" + "Shift Gear" + "Boost", designers may need to consider multiple "close" solutions and their preference score (see calculation example in Figure 8.8). Note that left and right hand/leg may be used simultaneously[2].

For actions such as "Shoot", "Drive", "Open Door", "Open Chest", hands were most preferred. Legs were most preferred for navigating the game avatar. For leg gestures, the ability to maintain good body balance was the participants' top priority. For example, participants commented that while performing certain 1-leg gestures (e.g., raising one leg up) they can easily lose body balance and thus may not be effective during fast gameplay. Meanwhile, for 2-leg gestures, fatigue was the participants' primary consideration, thus designers may try to avoid tiresome 2-leg gestures. In cases when legs are occupied, "Walk", "Run", and "Stealth Walking" actions could be optionally performed using two-hands on the sides of the body swinging back and forth, with the swinging velocity determining the different actions. For view-point changes, there was a strong consensus for using head movement as the primary input, but in cases when the head input was already assigned,"View point change" could

[2]Participants reported that left and right body parts (hand/leg) may be used simultaneously.

No.	Events	Suitability of body part						
		1-hand	2-hand	1-leg	2-leg	Head	Eye	Torso
1	**Accelerate**	2.75	2.75	4.75	1	1	1	1
2	**Boost**	6.75	3.5	1.75	1	1	1	1
3	**Break**	2.75	1.75	6.25	1	1	1	1
4	**Catch ball**	4	6.5	1	1	1	1	1.25
5	**Climb**	1	6.5	1	2.5	1	1	1
6	**Crawl**	3.5	2.5	1	2	1	1	1
7	**Dodge**	1	1	1	1	1	1	6.5
8	**Drive**	6.25	6.75	1	1	1	1	2.25
9	**Guard**	5.5	6.75	4	1	1	1	1
10	**Headbutt**	1	1	1	1	7	1	1
11	**Hide**	2	2.5	1.25	2.5	1.75	1	6.25
12	**Hitting baseball**	4.5	4	1.75	1	1	1	1
13	**Jump**	1.75	1	3	5.5	1	1	1
14	**Kick**	1.75	2.25	7	2.75	1	1	1
15	**Open chest**	5.25	4.25	3.25	2.5	1	2	1
16	**Open door**	6.25	3.25	4.75	1	1	2	1
17	**Open menu**	3.75	6.25	2	1.75	1.75	1	1
18	**Pick item**	6.5	5.25	1.75	1	1	2.25	1
19	**Punch**	7	6.5	1.75	2.25	1	1	1
20	**Push box**	6.25	6.25	3.25	1	2	1	1
21	**Racket hitting**	5.75	4	1.75	1	1	1	1
22	**Reload gun**	4.75	5.75	2.25	1	1	1	1
23	**Roll body**	2.5	2	1	2.5	1	1	1
24	**Roll object**	6	6.25	3.25	1	2	1	1
25	**Row boat**	5.25	6.25	1.75	1	1	1	1
26	**Run**	1	3.5	2	5.5	1	1	1
27	**Select menu item**	5.5	4	3	2	1	2.25	1
28	**Shake item**	6.5	6	1	1	1	1	1
29	**Shift gear**	6.5	2.25	3.25	1	1	1	1
30	**Shoot**	5.75	5	1.5	1	1.25	1	1
31	**Slash**	6.25	5	2.5	1	1.75	1	1
32	**Steal**	6.5	3.25	1	1	1	1	1
33	**Stealth walking**	1	1.25	1	5.5	1	1	1
34	**Stop walking**	6	2.75	3.75	1.5	1	1	1
35	**Throw ball**	6.25	2.75	1	1	1	1	1
36	**Use item**	5.75	5.5	2	1.75	2	1	1
37	**View point change**	3.5	1	2.25	1	6.5	5	1.75
38	**Walk**	1.25	3.5	2	5.5	1	1	1
39	**Zoom-in**	4.5	5.75	2.5	1	2.25	1	1
40	**Zoom-out**	4.25	5.25	2.5	1	2	1	1

Most suitable body part

Most suitable alternative body part

Figure 8.7: A set of suitable and alternative body parts (7 as most suitable).

also be optionally performed using eye movement or 1-hand movement as a cursor indicating the point of view. Eye movement was also promising for target acquisition tasks (e.g., “Pick Item”) but some participants mentioned that it could be difficult to control their eye movements. Torso gestures were most (and the only) preferred body part for torso actions (e.g., “Dodge”, “Hide”).

			Composite Score
1-hand Drive (6.25)	**1-hand Shift Gear** (6.5)	**1-leg Boost** (1.75)	**14.5**
	1-leg Shift Gear (3.25)	**1-hand Boost** (6.75)	**16.25**
	Other body parts (1)	**Other body parts** (1)	8.25
2-hand Drive (6.75)	**1-leg Shift Gear** (3.25)	**1-leg Boost** (1.75)	11.75
	Other body parts (1)	**Other body parts** (1)	8.75
Torso Drive (2.25)	**1-hand Shift Gear** (6.5)	**1-hand Boost** (6.75)	**15.5**
	1-hand Shift Gear (6.5)	**1-leg Boost** (1.75)	10.5
	1-leg Shift Gear (3.25)	**1-hand Boost** (6.75)	12.25
	1-leg Shift Gear (3.25)	**1-leg Boost** (1.75)	7.25
	Other body parts (1)	**Other body parts** (1)	4.25

Figure 8.8: Example of using the suitable and alternative body part from Figure 8.7 to determine possible combinations of "Drive"+"Shift Gear"+"Boost". Designers may choose among "close" score solutions (e.g., highlighted blue).

For system commands, symbolic gestures were preferred, e.g., "Zoom-in/out" action which could be performed using either a hand or a leg; a clenched fist can indicate "Zoom-in" and an open palm can indicate "Zoom-out" or, alternatively, raise a foot to 45 degrees to indicate "Zoom-out" and lower the foot for "Zoom-in". "Open menu" could be performed by swiping an index-finger down in the air, or making a camera gesture (imitating a photo-taking gesture by having the thumb and index finger of each hand resembling an L-shape sign, combining both hands to form a rectangle) or by dragging one foot from 12 o'clock to 6 o'clock or by performing a rotating movement with the head. "Select menu item" could be performed using a hand or the eyes to move the cursor, or by moving the legs to a different position (up, down, left, right) of the clock (three o'clock, six o'clock etc). The strategies of participants for system commands was that these commands should accommodate fast access during the middle of a intense gameplay.

8.6 Discussion (II)

Transferability between hand and leg. Our resulting set of suitable and alternative body parts posed a potentially useful, interesting design implication - the possible transferability between hand and leg gestures. When the hand(s) is occupied, the leg(s) tends to be the participants' preferred alternative, and vice versa. For example, gun-gesture using hand may be the most preferable gesture for "Shoot", but when the hand is occupied with other tasks, participants suggested using kick-gesture for “Shoot”. In a similar way, “Jump” can be performed alternatively by facing a palm up and raising the hand when the leg is occupied. Players could also simultaneously perform the “Use item” action by “double clicking” (tapping) one foot on the floor, instead of using hand(s). Together with “Drive” using driving-wheel gesture, “Break” and “Accelerate” can be simultaneously performed by stamping the foot to 12 o’clock or 6’clock. These transferability implications could prove handy for designers when designing simultaneous gestures. Furthermore, some participants suggested the possible gesture transferability between left and right hands in cases where one hand is already assigned. For example, one participant commented that although one hand is used for “Shoot”, another hand may be used for “Use Item” or “Reload” gesture. Similar comments were reported for left and right leg.

Limitation. In Experiment II, we were able to identify various game gestures using different body parts (e.g., “Walk” using 2-hands swinging across the body, or 2-legs moving like actual walking, 1-foot on the 12-o’clock position). Nevertheless, there is a need to further investigate the preferred combinations of gestures. Simply combining these gestures based on the designers’ intuitions may result in combined gestures that may not be anatomically comfortable or feasible. We addressed this issue in Experiment III.

8.7 Methodology (III)

In this experiment, we seek to develop a consensus set for simultaneous game gestures. Prior to this experiment, we conducted a pilot experiment where three participants were asked to elicit simultaneous gestures for a set of combined game events (e.g., shoot + reload). Nevertheless, although all our participants were veteran gamers, we found little consensus between their gestures. They also mentioned that the user-elicitation process was substantially difficult, as they commented that it was difficult to imagine possible combinations of gestures. Furthermore, they commented that the lack of any existing simultaneous gestures for interaction adds to the difficulty as there were few reference points.

Based on this challenge, we adapted the original user-elicitation method [78, 123] and introduced a novel choice-based approach for investigating suitable simultaneous gestures.

8.7.1 Choice-based elicitation approach

On the basis of human-centered design [63] and the original user-elicitation method [78, 123], we proposed a choice-based elicitation approach which is intended for cases when users do not have clear ideas or familiarity with the output space, possibly due to the novelty or the complexity of the requested output and thus may not effectively produce suitable gestures as claimed by Morris et al. [78]. In our case, since the notion of simultaneous gestures is relatively uncommon for users, with few reference points, we used a more suggestive, hinted approach (choice-based elicitation approach) to guide and assist our users to discover better, more suitable simultaneous gestures.

The main difference between a choice-based elicitation approach and a user-elicitation approach is a predefined-list of possible gestures. The predefined gesture list registers two data columns, one for possible gestures, and another column for game events, with the relationship as many-to-one, respectively. The predefined gesture list is intended to assist users to discover new or better possibilities when asked to perform certain gestures. In any case, users are also encouraged to define their own gestures and do not necessarily need to pick gestures from the predefined list when they preferred other options. In our case, the predefined gesture list was populated by results from our first two experiments.

The choice-based elicitation approach makes a basic assumption that in an unfamiliar scenario, users may lack knowledge of the design space, thus they may not always define effective gestures (as seen in Experiment I). The second assumption is that “recognition” may be better than “recall” in an unfamiliar elicitation scenario, i.e., given a list of choices (gestures), users will be able to better define/discover more suitable gestures (i.e., recognition), as opposed to only imagination from scratch (i.e., recall or original generation) which participants in our pilot study found to be relatively more difficult when the scenario is unfamiliar. The third assumption the choice-based approach made is that all participants should be experienced gamers in both general video games as well as full-body games, and understand the requirements of intense gameplay, so as to increase the likelihood of a suitable gesture set.

Corresponding to these assumptions, we hypothesize the benefits of choice-based elicitation approach as three-fold: first, it may allow participants to discover better, more suitable gestures from the predefined gesture list than they have originally come up with. Secondly, it may also allow participants to better understand the design space through examples, and allow them to become more "creative" in creating their own gestures. Thirdly, it may increase the likelihood of achieving a high level of consensus among participants. These benefits were investigated in this experiment in which we found positive results. Further details are described in the discussion section. In any case, one should also consider the possible threat of the choice-based elicitation approach which may prevent the definition of more novel gestures, i.e., participants may become "lazy" in defining their own gesture, and rely mainly on the predefined-list. We considered this threat in our study design.

8.7.2 Participants

Twelve veteran gamers (all males, M=22.4 years) were recruited. We selected primarily veteran gamers due to their better understanding of the needs and requirements, and of games in general (i.e., intended users). All reported to have experience with Kinect and motion game gestures. Ten participants were right-handed. They were each paid $10.

8.7.3 Events

We derived five combinations of common simultaneous game events or closely simultaneous events (transactional events[3]) from typical interactions in video games and full-body games, similar to Experiment I. The elicited events included the various combinations of the following game events: Shooting events (Shoot + Reload + Walk/Run/Stealth walking + Use First

[3]Transactional events refer to events that are executed quickly in sequence. Within transactional events, the prior event enables the execution of the next event, whereas simultaneous events do not pose such technical requirements. We used ⇢ to depict transactional events, while using + for simultaneous events.

Aid Kit + View point change), Racing events (Steer + Accelerate/Brake + Shift gear + Boost + View point change), Adventure/Role-playing events (Walk/Run/Stealth walking + Open door/Steal/Open chest/Pick Item/Push box/Slash + View point change), Fighting/Sport events (Punch + Kick + Guard + Dodge + View point change), System events (Open menu ⇢ Select menu item ⇢ Use item).

We used "/" (e.g., Walk/Run/Stealth Walking) to depict a category of events that can only occur one at a time, e.g., "Walk" cannot be executed at the same time as "Run". We assumed these events were interchangeable, so we asked participants to define simultaneous gestures that combine each of the following (e.g., Shoot + Walk, Shoot + Run, Shoot + Reload).

8.7.4 Procedure

At the start of the experiment, participants were asked to define gestures for the five combinations of game events (i.e., Shooting, Racing, Adventure, Fighting, System). The order of the five groups was asked in a randomized order across participants.

To reduce the cognitive burden on participants during the elicitation process, participants were instructed to first perform a set of two combined events (e.g., Shoot + Reload; Stealth walking + Steal). Then the number of combined events gradually increased by adding one more event (e.g., Shoot + Reload + Use First Aid Kit), until the combinations cover the full set of events (e.g., Shoot + Reload + Zoom scope in + Run + Use First Aid Kit + View point change). The gradually-increasing combined events were randomized. Although in reality, all these events may not (or very rarely) occur simultaneously (all at the same time), by asking participants to assume the most difficult case, we were able to detect and avoid any body part conflicts in the final combination of gestures elicited by participants. Our observations showed that participants consistently revised their gestures to be more simultaneous-friendly when the number of events increased.

Combined events were portrayed on a large display by displaying them separately, with a plus-sign "+" between target scenes along with the name of the event, e.g., "Shoot target scene" (event name) + "Reload target scene" (event name). By not mixing the target scenes, this separation minimizes a possible deterministic influence in the selection of gesture combinations. Most target scenes were sourced from Experiments I&II.

During the elicitation, to ensure the choice-based approach would not minimize the level of creativity of the participants, participants were asked to first do their best to perform the indicated combined gesture without the presence of a predefined gesture list. After the first attempt where participants performed each combined gesture or could not come up with a combined gesture, participants were exposed to a predefined list of possible gestures (populated from the results of Experiment I&II), where users could opt to refine their gestures based on the predefined list to produce their preferred simultaneous gestures. After their redefinition, to make sure that users' mix-and-match simultaneous gestures were anatomically and kinetically feasible and comfortable, we asked participants to again perform the actual simultaneous gestures quickly.

Similarly as in Experiment I&II, a think-aloud protocol was used where participants indicated the start and end of their performed gestures, the body parts used, and described the corresponding rationale. Other experimental place settings were similar to that used in Experiment 1. At the end of the experiment, we also asked the participants to rate the usefulness of the choice-based approach and performed semi-structured interviews. The experimental session took around 1.5-hours.

8.8 Results and analysis (III)

The agreement score, the resulting set of simultaneous game gestures and subjective assessment of the choice-based approach were analyzed and reported.

8.8.1 Agreement score

We used a similar approach to Experiment I to calculate the agreement score. Two gesture combinations were considered identical if each gesture in both combinations had similar trajectories and poses, and if the two combinations contained exactly the same set of gestures.

For example, for the agreement score of the scenario "Shooting", a total of 12 participants and their corresponding 12 sets of combined gestures were considered. This event was divided into 6 groups of identical combinations with sizes of 5, 2, 2, 1, 1, and 1. Therefore, the agreement score for the Shooting scenario was:

$$A_{shooting} = \left(\frac{5}{12}\right)^2 + \left(\frac{2}{12}\right)^2 + \left(\frac{2}{12}\right)^2 + \left(\frac{1}{12}\right)^2 + \left(\frac{1}{12}\right)^2 + \left(\frac{1}{12}\right)^2 = 0.25$$

Figure 8.9 shows the agreement score for each game event. The mean agreement score was 0.27, slightly lower than Wobbrock et al. [123](0.32 and 0.28).

There was greatest consensus in the fighting and racing gesture sets, with shooting and adventure gestures scoring almost as low on agreement as the system gestures.

Agreement score of each gesture (which is independent of the combination score) was also calculated (see the numbers in Figure 8.10 to Figure 8.14), where low agreement score in the overall combination was shown to be influenced by the disagreement in some game events.

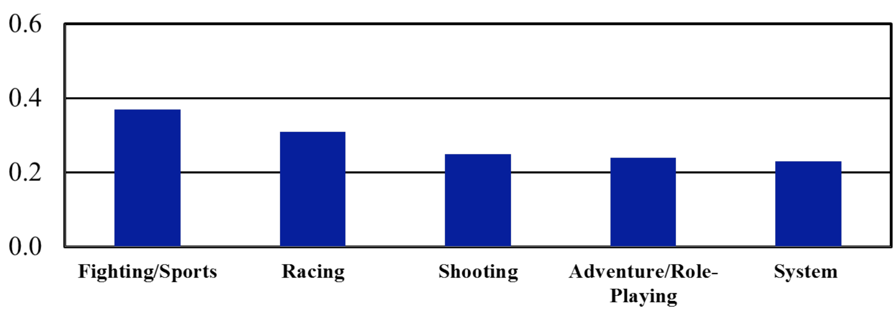


Figure 8.9: Experiment III - Overall agreement scores for the five combinations of game events sorted in descending order.

8.8.2 User-defined simultaneous gesture set

Of all the gestures performed, 95.3% of gestures originated from the predefined list - 61.1% of gestures were changed after seeing the proposed set and 34.2% were part of the predefined list although the gesture may have been invented by the participant without the knowledge of the set. Only 4.6% were newly created gestures elicited by participants. Although these gestures were newly created, participants reported that the predefined list provides a useful reference for imagining new gestures. All participants reported that without the gesture list, it would be difficult to imagine and design suitable gestures especially for simultaneous purposes with which they are not familiar. For example, participants said: "*The gesture list provides a good reference.*"; "*It's like a game, mixing and matching to find the best combinations.*"; "*The gesture list makes me much more creative*". This result clearly demonstrated the positive influence of this predefined list on the quality of the results.

Similar to past user-elicitation studies, we picked the largest group of combined gestures for our resulting set of simultaneous gestures (see Figure 8.10 to Figure 8.14). The priority of participants was simultaneous gestures that can be executed effortlessly and which are easy to learn/remember. Participants also preferred to reuse similar gestures across game events in different game groups.

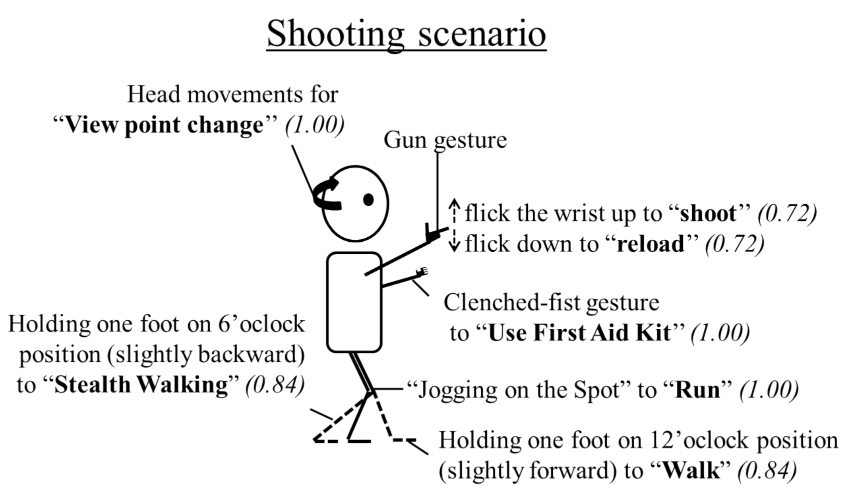


Figure 8.10: User-defined simultaneous gesture set for shooting scenario. Note: The numbers depict the agreement score of each gesture, which are independent of the overall combination's agreement score.

Shooting- "Shoot" action was most preferably performed using one-hand imitating a gun-gesture and then flicking the gesture up (flicking the wrist up). Simultaneously, “Reload” action was most preferably performed by flicking that one-hand gun-gesture down. “Use First Aid Kit” was most preferably performed by performing a clenched-fist gesture using the remaining available hand. Head movement was most preferred for “View point change”. For game avatar navigation, “Walk” was most preferably performed by holding one foot in the 12 o’clock position (i.e., pushing the foot slightly forward); “Stealth Walking” was most preferably performed by holding one foot in the 6 o’clock position (i.e., pushing the foot slightly backward); “Run” was most preferably performed by performing an actual run action, by jogging on the spot repeatedly.

Shooting scenario achieved a mean agreement score of 0.25. The differences came from minor disagreement of various gestures. For “Shoot”, alternatively, some preferred a two-hand gun gesture (one hand as gun-gesture, while another hand holds the gun-gesture). For “Reload”, some preferred using the other hand hitting the bottom of the gun-gesture to “Reload”. For “Walk”, some preferred to perform actual walk action. For “Stealth Walking”, some preferred a slow-motion walking action.

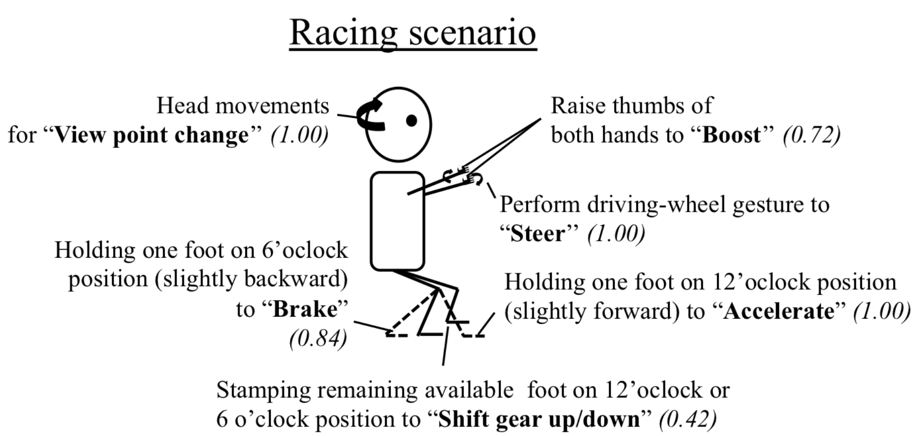


Figure 8.11: User-defined simultaneous gesture set for racing scenario. The gesture set assumed a user in a sitting posture.

Racing- “Steer” action was most preferably performed using two-hands to perform a driving-wheel gesture. "Accelerate" and "Brake" were most preferably performed by holding the foot to 12 o'clock and 6 o'clock respectively. While both hands are performing a driving-wheel gesture ("Steer"), most participants preferred to raise the thumbs of both hands to indicate "Boost". Head movement was most preferred for performing "View point change" action. For "Shift gear" action, the most preferable gesture was to stamp the remaining available foot to 12 o'clock and 6 o'clock position for shifting the gears up and down, respectively.

The Racing scenario achieved a mean agreement score of 0.31. The main differences came from the definition of "Boost", "Shift gear up/down" and "Brake" gestures. Alternatively, participants preferred pushing the driving-wheel gesture forward to "Boost". For "Shift gear up/down", some participants preferred stamping the foot on 3 o'clock and 9 o'clock position for shifting the gears up and down, respectively. For "Brake", some participants preferred to hold one foot on the center of the clock instead to the 6 o'clock.

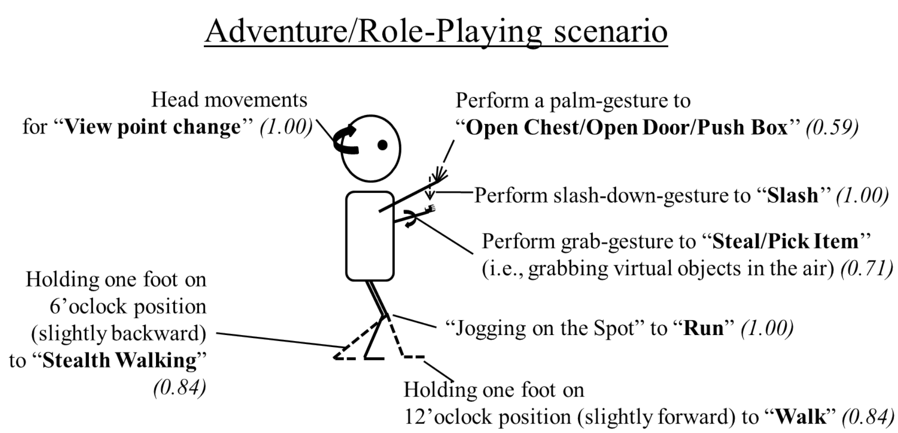


Figure 8.12: User-defined simultaneous gesture set for adventure scenario. Participants reused navigation gestures (Run, Walk, Stealth Walking) from the "Shooting" scenario.

Adventure/Role-Playing-For game avatar navigation, participants preferred similar gestures to navigation in Shooting events. Jump was most preferably performed by raising one leg slightly. Open door/Open chest/Push box/Steal/Pick Item/Slash events were most preferably performed using hand(s). For "Open Chest", "Open Door" and "Push box" actions, the most preferable gesture was to perform a palm gesture using 1-hand; for a "Slash" action, the most preferable gesture was to use one arm performing a slash-gesture; for "Steal" and "Pick Item" action, the most preferable gesture was to use one hand performing a grab-gesture. Head movement was most preferred to perform a "View point change" action.

The Adventure/Role-playing scenario achieved an agreement score of 0.24. The main differences came from whether participants preferred a hand or leg to perform certain actions. For example, to "Open Chest" or "Open Door", some participants performed a kick-gesture. Participants also showed different choices whether to use palm-gesture/punch-gesture/grab-gesture to "Open Chest", "Open Door", "Push Box", "Steal".

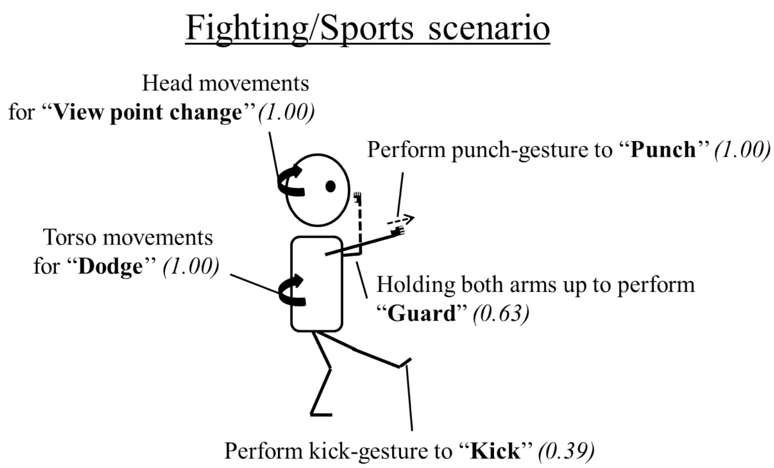


Figure 8.13: User-defined simultaneous gesture set for fighting scenario. Gestures performed in this group were straightforward and mostly imitating real-world gestures due to the few body part conflicts.

Fighting/sports-Events in this group were performed by fairly predictable and straightforward gestures, mostly imitating real-world actions due to the few body part conflicts, i.e., "Punch" was preferred by performing a punching gesture, with either hand. "Guard" action was preferred by holding both arms up on guard in front of the user's face. "Kick" action was preferred by performing a kick gesture. "Dodge" action was most preferably performed by moving the torso left and right accordingly. Head movement was most preferred for performing "View point change" action.

The Fighting/sports scenario achieved a relatively high agreement score of 0.37, due to obvious counterparts in the real-world. The main differences came from the definition of "Kick" gesture. Participants performed multiple variations of Kick such as straight-kick (most common), sideward-kick, knee-kick (using the knee to kick), feet-kick (using feet movement). For "Guard", some preferred one-handed guard.

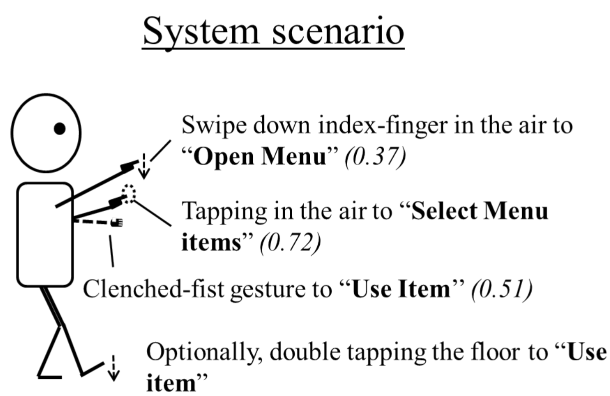


Figure 8.14: User-defined simultaneous gesture set for system scenario. Unlike other simultaneous scenarios, these events were treated as transactional events (events in quick sequence) thus the same body parts were allowed to be used for different events.

System-"Open menu" command was most preferably performed by using the index-finger of either hand, and swiping down in the air to open a menu. "Select menu items" command was most performed by tapping in the air using the index finger and the "Use item" command was most preferably performed by performing a clenched-fist gesture. Nearly-equal in preference (6 participants), the "Use item" command can be optionally performed by "double clicking" (double tapping) one foot on the floor.

The System scenario achieved an relatively low agreement score of 0.23. "Open menu" could be alternatively performed by other symbolic gestures such as an opening-window gesture (virtually opening a window in the air), pushing-button gesture (pushing virtual button in the air using palm), and camera gesture. One could alternatively perform the "select menu item" using legs or eyes as the cursor. Given the relatively low agreement score of system events, it may be helpful to allow end-user customization on these events.

8.8.3 Subjective assessment of choice-based elicitation approach

Participants were asked to rate the choice-based approach based on the Likert 7-scale rating (7 as strongly agree). Overall, participant responses to the choice-based approach were encouraging. The mean value is "Using the choice-based approach, the simultaneous gestures I performed were a good match for its intended purpose" (6.34); "The choice-based approach guides me to more easily discover better gestures" (6.23); "I would prefer using the choice-based approach in an unfamiliar elicitation scenario" (6.10); "The choice-based approach makes me more creative in creating new gestures" (5.32).

8.9 Discussion (III)

We discuss the (i) effectiveness and challenges of the choice-based elicitation approach, (ii) informal comparison between our defined gesture set with existing Kinect games, and (iii) implications for full-body recognition technology.

8.9.1 Effectiveness and challenges of the choice-based approach

Based on our subjective assessment and semi-structured interviews, we confirmed our hypothesis regarding the benefit of the choice-based elicitation approach. Participants commented on the usefulness of the choice-based approach in two primary ways: first, it allows participants to discover more efficient, suitable gestures (better than they originally imagined/elicited); second, it provides participants with a reference point from which they can further develop their own gestures. For example, participants commented that they felt that they become more "creative" after they saw the gesture list, and could subsequently come up with several creative gestures that did not exist in the original gesture list.

The choice-based elicitation approach also allows us to achieve a high consensus - agreement score of 0.27 - which is considered relatively reasonable when compared with past works Wobbrock et al. [123]. We also compared the agreement score between the initial gesture set that participants came up with (before showing the predefined list) and with the final gesture set. The average agreement score for the final gesture set (0.27) was around two times higher than that of the initial gesture set (0.11). Overall, the choice-based elicitation approach showed a positive effect as we had hypothesized.

In any case, our study raised several issues for choice-based elicitation. *First*, participants commented that excessive choices would add cognitive burdens to participants, and suggested that approximately three to six choices would be the optimal number of choices. Too few or many choices could reduce the effectiveness of the choice-based approach, thus there is a need to choose which appropriate gestures are to be listed. We regard such a selection process to be at the core of any choice-based elicitation approach, and it requires solid, objective criteria if designers are to ensure that the intended quality of the predefined gesture list is as free of deterministic bias as possible. In our case, we selected the top-6 most-elicited (preferred) gestures from Experiment I&II to be listed. In other cases, the choices can depend on the criteria set by designers (e.g., top-X most comfortable). To prevent any possible ordering effect, it is also important to randomize the order of choices in the predefined-list. *Second*, there might be a possibility of a choice-based approach to reduce participants' level of creative imagination and confidence. It is thus important to not expose the predefined-list before participants elicit their first trial gesture or give up on elicitation due to lack of ideas. Participants should also be encouraged to define their original gesture when possible. *Third*, the approach requires a predefined gesture list whose creation can be time-consuming and it requires expert-participants who may not always be available. Based on these issues, there is a need to further investigate and compare the effectiveness of the choice-based elicitation approach with the traditional user-elicitation approach.

8.9.2 Comparing our gesture set with existing games

We compared our user-defined gesture sets with some existing Kinect games. We found that many Kinect games do not accommodate the simultaneous use of gestures. For example, existing First Person Shooting Kinect games (e.g., Gunstringer, Ghost Recon, BlackWater) primarily used hand(s) as the input modality, such as raising the left hand up to indicate “Jump”, using a punch-gesture for “Shoot”, using a touching-elbow gesture for “Reload”, and a hand to control “View point change”. On the other hand, locomotion is automated, probably to reduce physical complexity. Although these hand gestures may be comfortable to execute, they do not facilitate the simultaneous use of game gestures. We proposed in our user-defined gesture set that “View point change” can be controlled by head movements, or alternatively by eye movements; “Running” can be achieved by jogging on the spot, “Reload” can be done by flicking the wrist down, so as to allow “Shoot” and “Reload” to be achieved simultaneously, instead of the touching-elbow gesture which would interrupt the “Shoot” action.

Existing racing Kinect games (e.g., Kinect Joy Ride, Forza 4) also mainly used a hand for interaction, such as performing the driving-wheel gesture to “Steer”, and pulling back the wheel gesture to “Break”. We proposed in our gesture set that leg modality can be leveraged for performing “Break”, “Accelerate” and “Shifting gear” which can better accommodate the simultaneous use of gestures.

We found that several existing Kinect games limit the player’s degree of control. For example, some games (e.g., Forza 4, Kinect Joy Ride, BlackWater) attempt to reduce physical complexity by removing and automating certain actions (e.g., Walking, Accelerate). However, although automating certain actions might remove physical complexity, all of our participants prefer the ability to control their movements (stop walking, walking, running). This is because the sense of control is important for achieving high immersion in the game. For example, one participant stated: “*Many existing Kinect racing games are boring because there is no realism. I cannot control my speed. I cannot feel immersed in racing; I am simply* *stretching my arm.*"

Some Kinect games used more "efficient" gestures (punch-gesture for "Shoot") over the natural gesture (gun-gesture for "Shoot") but this may not always be beneficial. For example, although the punch-gesture may perform more precise shooting, this might disrupt immersion. One participant stated: "*I would prefer gun-gesture over any other gesture for shooting, if not, I would never able to become immersed in the game*". The tradeoff between efficiency and naturalness should be carefully balanced.

For the System scenario, such as open menu or selecting menu items, existing Kinect games often employed a simple hand-pointing method (i.e., using hand as cursor). Our user-defined gesture set informs that users prefer a symbolic gesture for opening a menu (e.g., swiping down index-finger in the air), which can be less time-consuming compared to waving the hand for a long time to select menu items. Participants' priorities are to facilitate easy and fast access to system commands while in the middle of an intense gameplay.

For the Adventure and Sports scenario, we did not find much difference between existing games and our gesture set. The gestures are mostly pantomimic (one-to-one correspondence between the gesture and game events), e.g., punch-gesture for punching. This is probably because adventure and sport game events have obvious real-world counterparts in which there are few conflicts in body-part usage.

8.9.3 Implications for full-body recognition technology

Many of the simultaneous gestures had strong implications for full-body recognition technology in highly interactive game scenarios. For example, with the need of simultaneous gestures in highly interactive scenarios, the recognizer should able to support recognizing (and distinguishing) simultaneous gestures at one time.

In addition, our resulting gesture set also indicates that the recognizer should be able to accommodate any slight variations of the similar gestures. For example, to perform "Kick", players may perform kick gesture at different heights and angles, with each gesture posing slight natural variations. Another example is "Shoot" where some participants may use one-hand gun gesture, while others may prefer holding the gun-gesture using two-hands.

With detailed gestures such as "gun-gesture", "index-finger swipe" gesture or "clenched-fist" gesture, the recognizer should be able to distinguish gestures to the details of the hand posture, number of fingers used, and which finger was used. For combinations of events with a low agreement score (e.g., system scenario), end-user customization should be enabled.

Considering the features of current technologies (e.g., Kinect), we believe it is technologically feasible to meet these needs but the main challenge lies in detection and analysis algorithms (i.e., image processing). Some studies regarding natural variations [38] and finger detection [69] have already been started and showed promising results, but less has been done on the detection of simultaneous gestures. This reflects the need for concurrent gestures detection and analysis algorithms, particularly for video games and other fast-paced interaction applications.

8.9.4 Design principles and guidelines

We summarized our findings into generalized design guidelines.

- **Guideline 1: Prioritize events.** One difficult decision designers may encounter is regarding the assignment of gestures. For example, it can be difficult to decide whether to use the most suitable body part or the alternative body part for one event. Our studies show that participants define gestures according to priority. Participants commented that they would used the most suitable body part for crucial/frequently-used game events (e.g., hand for Shoot), while using symbolic/abstract gestures for less frequently-used game events (e.g., leg for open menu). The participants' general strategy is to minimize their effort to learn and remember combined gestures and also maintain the maximum degree of naturalness possible. Thus it is important for designers to prioritize which events are more significant than others.

- **Guideline 2: Prioritize immersion over playability.** Several existing Kinect games simplified the interaction by removing certain actions such as "Walk" or "Accelerate". However, this can have an adverse effect on player's immersion, as one of our participant stated, "*Game gestures should be the same as real-world actions wherever possible, if not, I can easily get distracted and I cannot feel immersed*". The lack of correspondence between the real-world and virtual-world can disrupt player's immersion and presence in the game. All our participants mentioned that immersion is among their top reasons in playing full-body game interactions instead of using traditional controllers (e.g., a keyboard).

- **Guideline 3: Use the hand moderately.** Although the hand is unsurprisingly the most preferred body input modality, it should be used moderately considering possible simultaneous use with other gestures. One should also consider the distribution of fatigue as one participant mentioned "*Using only hand is super tiring, I prefer my fatigue to be equally distributed across my body, which is also more fun.*"

- **Guideline 4: Exploit transferability between leg and hand and right and left body parts.** We found the possible transferability between hand and leg gestures, and between left and right body parts (e.g., left hand and right hand). This suggests that when certain body parts are occupied, designers may exploit this transferability when designing gestures. For example, when "Shift Gear" cannot be performed by a hand, the event then can be performed by a leg. In a similar manner, when the left hand is occupied with "Shoot", then one can use the right hand to perform other hand events (e.g., Use Item).

- **Guideline 5: Accommodate high tolerance for recognizing gestures.** We found that although all participants prefer a similar gesture for one event, they execute slightly differently in velocity and displacement based on their individual preferences. For example, for "Kick", participants kicked at different heights, angles and velocities. Thus recognition systems should tolerate such natural variations. Providing customization may help address this issue, e.g., allowing participants to customize different velocity/height for different kinds of "Kick".

- **Guideline 6: Gesture reuse.** Gesture reuse is important to assist the learnability of users. For example, our study found that reversible gestures were preferred by participants for reversible actions (Walk/Stealth Walking, Accelerate/Brake, Shift gear up/down). Similarly, participants preferred the same gesture for "Walk" in the Shooting scenario, and for "Accelerate" in the Racing scenario. Given this information, the idea of reusing gestures and designing reversible gestures has support.

- **Guideline 7: Design multiple gestures for one event, when needed.** We found individual differences between gamers and non-gamers. To accommodate differences in game expertise (novice vs. expert), one approach is to design more than one gesture for one event just as in desktop-based system. For example, designers can design two gestures for "Jump" - actual jump gesture for novice players and a more symbolic jump gesture (e.g., raising one leg slightly above the floor) for expert players.

- **Guideline 8: Reducing fatigue by a small amount can have great impact.** A small reduction in fatigue can have great impact in lengthy gameplay. For example, three participants mentioned that flicking the wrist up to "Shoot" is very different from flicking the lower arm in a long gameplay. Gestures should be designed to optimize the required efforts in lengthy gameplay.

- **Guideline 9: Design kinetically feasible combined gestures.** Designers should be careful to design kinetically-feasible combined gestures. Participants preferred simultaneous gestures that can be executed with ease and efficiency. Uncomfortable posture/gestures should be avoided (e.g., moving the head left while moving the torso right may cause injury). This suggests that full-body game designers should work with doctors who are specialized in motor control and coordination.

Chapter 9

STUDY 7: EMPIRICAL STUDY IN FULL-BODY GAMES (II)

This study aims to further evaluate the effect of player differences in full-body game engagement, concerning the moderator dimension of the framework (see Figure 9.1).

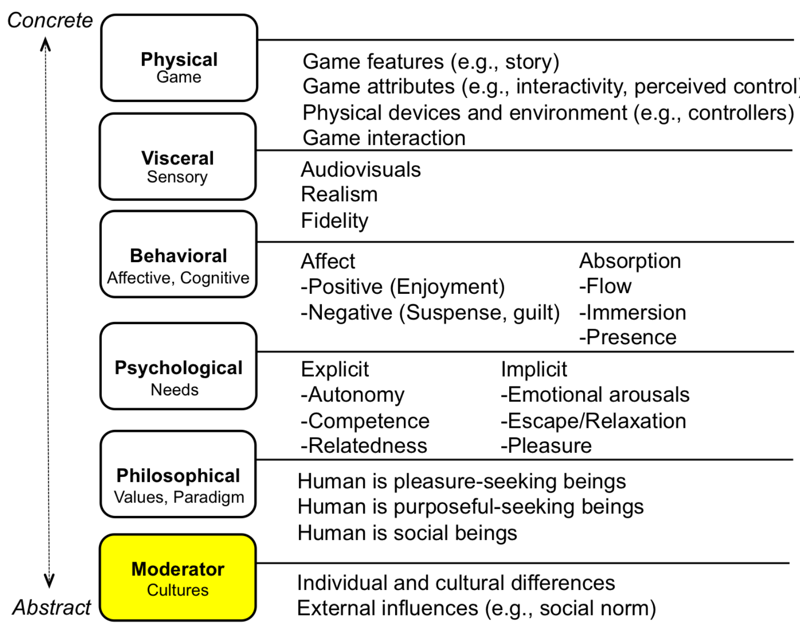


Figure 9.1: Study 7: Evaluating player differences in full-body games

Full-body motion gestures enable realistic and intuitive input into video games. However, despite their intuitiveness, players may not always prefer full-body gestures for gameplay. Indeed, some players have reported that they did not enjoy full-body-based games even though they seem reasonably usable and natural (e.g., [1]). Without a clear understanding of player differences and how each different player envisages to enhance their gaming experience, it is difficult for designers to develop engaging full-body game interfaces.

Figure 9.2: Three human factors possibly affecting player engagement: player's *motivation to succeed*; player's *motivation to move*; and player's *gaming expertise*.

Several studies regarding how body movement influences players during gameplay have been conducted [12, 79, 43, 83, 88]. Yet, the different kinds of players and how they engage/disengage in full-body games remained unexplored. For example, in which context will players engage/disengage full-body games? In full-body game interaction, who are the primary targeted users? If any? Are full-body games reserved for casual players only? How can designers better support the wide variety of players? Understanding the rationale behind players' preferences could enable designers to develop more enjoyable and effective full-body game interactions.

To investigate these questions, adopting a user-typing approach [6, 16], we explore player differences and their preferences in full-body game interaction. Specifically, we hypothesized three human factors that influence player's engagement (Figure 9.2). To explore the hypotheses, we conducted a formal experiment with 16 participants. The results revealed a significant correlation and main effect of the three factors on different subscales of player engagement. Further analysis also revealed three important game properties that associates with players' preferences: level of cognitive challenge, level of physical challenge, and level of realistic interactions.

9.1 Methodology

9.1.1 Hypothesis

Success Motivation Hypothesis. This hypothesis was formulated based on the findings of [83, 88] and our pilot study. This hypothesis predicts that when players are motivated to win/succeed or to achieve certain things in the games (achievers), they are motivated to search for the most efficient way to accomplish tasks. When a full-body game is competitive and offers efficient ways to interact, achievers will likely engage in the game. Vice versa, the lack of perceived competitiveness or inefficient control will likely hinder the (engagement and) experience of achievers. On the other hand, when players play the game solely to relax or to enjoy it with friends/family (casual players), the players are likely to enjoy and engage in full-body gestures, as they provide high levels of social and affective enjoyment [12] and relaxation. In addition, overly difficult challenges may not be perceived as fun/relaxing for casual players.

Movement Motivation Hypothesis. Player motivation to move refers to the player's general personality to enjoy or avoid moving. Two major groups of players can be classified: movers and non-movers. "Movers" refer to players who enjoy physical activities in real-world (sports or outdoor physical activity) and thus possibly transfer this motivation to digital games as well [34]. On the other hand, "non-movers" refers to players who prefer to avoid unnecessary movements, as they may consider full-body games to be tiring and cumbersome. The hypothesis predicts that movers are likely to engage in full-body interactions while non-movers possess a tendency to avoid full-body interaction when possible. According to [26, 88], the hypothesis also predicts that in full-body games, non-movers are likely to achieve higher engagement in full-body games that they perceive as playful rather than in games that they perceive as serious and competitive.

Expertise Hypothesis. Game expertise refers primarily to experience in the context of gaming frequency and players can be classified as gamers and non-gamers. There are two subconstructs of gaming expertise that may affect player's experience in full-body interaction.

First, given high game expertise, gamers may develop an affinity (bias) toward traditional inputs such as a gamepad controller or a mouse/keyboard interface according to the mere-exposure theory [126]. Second, since there is a high tendency for these gamers to exhibit longer sessions of gameplay compared to non-gamers, gamers may tend to avoid full-body interaction as they anticipate that the experience will be tiring over lengthy gameplay. Based on these two subconstructs, the hypothesis predicts that gamers might have some tendency to disengage full-body gestures when the interaction does not meet their expectations. On the other hand, non-gamers may be more easily engaged in full-body game interaction as it is possibly perceived as easy to learn and master, and natural and intuitive [64].

9.1.2 Selecting games

Before the study, several full-body games were explored based on several suitability criteria. The games should first be of high quality and thus pose little or no usability issues. Usability issues were generally understood to be possible recognizer issues or requiring awkward combinations of motion gestures. The high quality of games was quantified based on players' review scores from various gaming sites. Second, in order to keep the whole test setting within a reasonable time-frame, the games should also allow players to experience and enjoy the whole process of the game without necessarily spending overly long sessions of playtime. Third, the selected games should vary in their interaction paradigms thus allowing us to investigate how different kinds of players interact with different kinds of games. Three games were selected based on our criteria- *Virtual Tennis 4*, *Forza 4*, and *London Olympics 2012*. They varied in their relative perceived competitiveness, physical challenge, cognitive challenge, movement paradigm, level of realistic interaction, and hence provided helpful experimental tools to explore the hypotheses from different game types. Table 9.1 shows the relative differences.

Virtual Tennis 4 (VT). VT is a tennis-sports game. VT supports full-body input without any reported usability issues. The game is played in the first-person view. The task is to role-play as a professional tennis player and compete with other A.I. controlled players. During

Games	Perceived Competitiveness	Realistic Interaction	Movement paradigm	Physical challenge	Cognitive Challenge
Virtual tennis 4	High	High (like real-tennis)	Frequent and large axes of movement	High	High (required highly coordinated, frequent movements with the ball, arm, and the torso. Players also need to predict the ball movement from opponents.)
Forza 4	Medium	Low (players only required to steer left/right without having to brake or accel)	Infrequent and small axes of movement	Low	Medium (required eye attention on the road)
London Olympics 2012	Medium	Medium (partially realistic)	Frequent and medium axes of movement	Medium	Medium (required coordinated movements in some intervals)

Table 9.1: Overview of three games selected for study. They varied in their level of perceived competitiveness, level of realistic interaction, movement paradigm, level of physical challenge and level of cognitive challenge. They all provide a satisfactorily level of usability, i.e., adequate level of mimicry of movements and proprioceptive feedback. Realistic interaction refers generally to level of realism in controls and interactions, but not graphic-wise. Axes of movement refer generally to kinematics displacement. Physical challenges refer generally to the amount of body effort. Cognitive challenges refer to the amount of coordinating processing between perceptual processor, cognitive processor, and motor processor.

the gameplay, players are required to swiftly move left and right by swaying the whole body, and to swing the whole arm to the left or right accurately and with good timing in order to drive the ball back to the opponent.

Forza 4 (FZ). FZ is a racing game. FZ supports full-body input without any reported usability issues. The game can be played in the first-person or third-person view depending on the players' preference. The task is to complete a certain set of races. Players are only required to perform a driving-wheel gesture and to steer left and right to finish the race (see Figure 9.3).

London Olympics (LO) LO is a party game which is meant to be played at social gatherings. The task is to compete in different series of Olympic activities, e.g., running, hurdling, and javelin throwing. For the task, players were required to perform actual jumps according to timing and accuracy and to swing both their lower arms outward and inward repeatedly as frequently as possible (e.g., hurdle).

9.1.3 Design

Participants were tasked with playing three TV games using Kinect full-body gestures in a within-subject design. The sequence of the three games was randomized and counterbalanced using Latin Squares. Participants were instructed to complete a series of tasks for each game. To minimize any possible effect of difficulty factor, we ensured that all games were played at the same difficulty level and with similar game tasks across participants. To also ensure that there was no effect from possible visibility issues, all participants stood/sat approximately 1.8 meters away from the large display. During the gameplay, for our participants to fully enjoy and experience the games, we kept the testing environment away from possible interference, e.g., experimenters did not interrupt or ask any questions during the gameplay. The experiment was video and audio recorded for later analysis. After each game was tested, participants were asked to complete a self-report questionnaire measuring their preference and engagement. To collect qualitative data, we conducted an interview afterward.

Players' contextual info was collected using predefined forms comprising a series of questions. Gaming expertise was rated on a 5-point Likert scale with 5 representing high gaming expertise in a series of three items, i.e., gaming frequency per week, gaming hours per game per time, and gaming years. Gaming expertise was collected prior to the actual experiment

Figure 9.3: Participants play three games using Microsoft Kinect controller. *Virtual Tennis 4* (left), *Forza 4* (mid), *London Olympics 2012* (right).

session.

Players' general motivation to move was rated on a 5-point Likert scale with 5 as "strongly agree". It was measured using a series of six items, e.g., "I enjoy moving with my body", "I do not mind moving my body during gameplay". This information was collected prior to the experiment.

As for "motivation to succeed", we considered this as dependent upon different games. Thus, after each game was played, prior to the questionnaire session, players were asked a series of five items rated on a 5-point Likert scale with 5 as "strongly agree", e.g., "I put a lot of effort into the game", "I got easily stressed during the gameplay", "I play to relax my body and mind".

9.1.4 Participants

16 university students (3 females, M=21.75 years) were recruited. They were all from Computer Science backgrounds.

9.1.5 Apparatus

Microsoft Xbox 360 and Microsoft Kinect were used connecting with a SHARP Aquos 60-inch flat screen LCD display hung vertically. Other equipment and software included a Panasonic HDC-TM45 1920x1080 video camera, the three Kinect games, and related questionnaire forms.

9.1.6 Procedure

All participants were first informed about the aim and procedure of the study. They were then asked to sign consent a form regarding possible physical injury and to fill in demographic info (e.g., gaming expertise, general motivation to move). Before the gaming session, the three games were presented for 5-minutes and could be trialed until participants become familiar with the input and game mechanics, and were able to play by themselves. Then participants were assigned to play each game in randomized order. After playing each game, a questionnaire session (measuring players' engagement and motivation to succeed) and semi-structured interviews were conducted. Each experimental session of each game took around 15 minutes with a 5-minute break between. The whole experiment took around 1 hour per participant.

9.1.7 Metrics

The measurement of gaming experience involved largely self-reported measures, namely questionnaires. For our study, we used the Game Experience Questionnaire (GEQ) core questionnaire module [54]. The questionnaire measured seven dimensions: Immersion (Imm), Flow (Flo), Competence (Com), Tension (Ten), Challenge (Cha), Positive Affect (PoA) and Negative Affect (NoA). Each item was measured in a Likert 7-scale response with 7 as strongly agree. To keep the questionnaire within a reasonable timeframe, the social module and post-game module of GEQ were discarded. Three relevant items were added to the questionnaire:"I prefer to use Kinect for this game" (Prf), "I feel difficult when playing Kinect" (Dif), and "I feel fatigued using Kinect" (Fat).

9.2 Results and analysis

9.2.1 Demographic information

User demographic info scores (gaming expertise, motivation to move, motivation to succeed) were aggregated into an average score with high internal reliability of Conbrach α ranging from 0.86 to 0.92. We did not find any significant correlation between player types, e.g., no correlation between gamer participants and achiever participants.

We used cluster analysis to classify possible numbers of player types according to the participants' given demographic info. For each player's dimensions (gaming expertise, motivation to move, motivation to succeed), two primary groups of players can be classified: gamers vs. non-gamers, movers vs. non-movers, and achievers vs. casual players. Non-gamers scored an average of 1-2 in the scale of gaming expertise while the rest are gamers. Non-movers scored an average of 1-2.5 in the scale of motivation to move. Casual players scored an average of 1-2.5 in the scale of motivation to succeed.

For Kinect experiences, seven reported to have no experience in Kinect or full-body game interaction. Another eight reported one to three times, and one reported four times. None had ever experienced any of the three games.

There were three female participants in the study. All female participants were non-gamers. During the study, we did not observe any gender-specific differences between the female participants and the male non-gamer participants.

Figure 9.4 summarizes the distribution of participant demographic information.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16
Age	21	21	23	21	21	22	22	20	23	24	19	19	22	22	25	23
Gender	M	M	M	M	M	F	M	M	M	M	F	F	M	M	M	M
Kinect Exposure	No	Yes	No	Yes	Yes	Yes	No	Yes	Yes	No	No	No	Yes	No	Yes	Yes
Movers vs. Non-movers	M	M	M	NM	NM	M	M	NM	M	NM	M	NM	M	NM	M	M
Gamers vs. Non-gamers	G	G	G	G	NG	NG	G	NG	G	NG	NG	NG	G	NG	G	G
Achievers vs. Casual players (for each game)	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO	VT FZ LO
	A A C	C C C	A A C	A C A	C C C	A C C	A C C	A A A	A A A	A A A	A A A	C A C	A A C	A A A	A A C	C C C

Figure 9.4: Distribution of participants' demographic info. VT(Virtual Tennis); FZ(Forza); LO(London Olympics). We did not find any significant correlation between player types (e.g., achievers vs. gamers).

9.2.2 Quantitative Results

Questionnaire scores for each item (Immersion, Flow, Competence, Tension, Challenge, Positive Affect, Negative Affect) were aggregated into average scores since all scales produced high levels of internal reliability with Conbrach α ranging from 0.82 to 0.97. In addition to the three additional items (preference, difficulty, fatigue), total of ten scales were measured.

To analyze any correlation or significant main effects, Pearson correlation analysis and MANOVA were conducted.

Pearson correlation data

To analyze a possible correlation of Motivation to move, Motivation to succeed, Gaming expertise vs. Immersion, Flow, etc., we conducted the Pearson correlation analysis which resulted in the Pearson Correlation Matrix (see Figure 9.5). There were significant correlations for all three factors in different subscales of engagement suggesting their general association with player engagement.

As shown in Figure 9.5, across different games, there were significant correlations ($p<0.05$) between Motivation to move and Immersion, Competence, Tension, and Preference; significant correlations between Gaming expertise and Immersion ($p<0.01$), Flow ($p<0.01$), Positive Affect ($p<0.05$), Negative Affect ($p<0.01$), Preference ($p<0.05$), and Fatigue ($p<0.05$); and significant correlations between Motivation to succeed and Immersion ($p<0.05$), Flow ($p<0.01$), Competence ($p<0.05$), Positive Affect ($p<0.05$), Negative Affect ($p<0.001$), Difficulty ($p<0.001$), and Fatigue ($p<0.05$).

	Imm			Flo			Com			Ten			Cha		
Games	VT	FZ	LO	VT	FZ	LO	VT	FZ	LO	VT	FZ	LO	VT	FZ	LO
Motivation to move	**0.53***	0.37	0.23	0.37	0.21	0.13	**0.50***	**0.56***	0.24	**0.54***	0.22	0.17	0.40	0.08	0.26
Gaming expertise	0.34	0.10	**0.68****	0.38	0.01	**0.68****	0.43	0.05	0.45	0.11	0.48	0.13	0.31	0.07	0.33
Motivation to succeed	**0.54***	-0.34	0.007	**0.69****	-0.39	0.004	**0.57***	-0.06	0.46	0.30	-0.27	0.38	0.44	0.05	0.35

	PoA			NoA			Prf			Dif			Fat		
Games	VT	FZ	LO	VT	FZ	LO	VT	FZ	LO	VT	FZ	LO	VT	FZ	LO
Motivation to move	0.44	0.27	0.17	0.36	0.32	0.05	0.39	**0.57***	0.05	0.28	0.05	0.17	0.22	0.16	-0.42
Gaming expertise	0.45	0.19	**0.56***	0.35	0.005	**0.62****	0.35	0.18	**0.61***	0.31	0.18	0.47	**0.50***	-0.12	0.25
Motivation to succeed	**0.54***	-0.36	0.04	**0.79*****	-0.03	0.13	0.43	-0.36	-0.04	**0.78*****	-0.28	-0.07	**0.55***	-0.20	0.08

Figure 9.5: Pearson correlation matrix of Motivation to move, Gaming expertise, Motivation to succeed vs. seven subscales of GEQ. Three additional items are Preference (Prf), Difficulty (Dif), and Fatigue (Fat). (2-tailed Sig.: *=$p<0.05$; **=$p<0.01$; ***=$p<0.001$)

MANOVA

To analyze possible effects of the three independent variables (Movers vs. Non-movers, Achievers vs. Casual Players, Gamers vs. Non-gamers) on several dependent variables (e.g., Immersion, Flow), a three-way multivariate analysis of variance (MANOVA) was conducted.

Comparing movers and non-mover participants, there were significant differences in Immersion ($F_{1,14}$=5.052, $p<0.05$), Flow ($F_{1,14}$=4.676, $p<0.05$), and PoA ($F_{1,14}$=6.68, $p<0.05$)

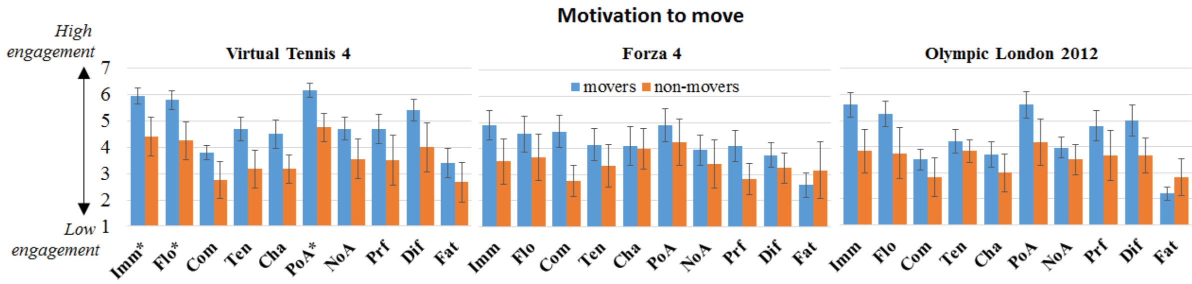


Figure 9.6: Motivation to move - a significant difference between movers and non-movers in Imm, Flo, and PoA. The difference was only found in Virtual Tennis. 7 represents high engagement; NoA, Dif, and Fat were reversed (e.g., 7 in Fat represents low fatigue). (2-tailed Sig.: *=p<0.05; **=p<0.01; ***=p<0.001)

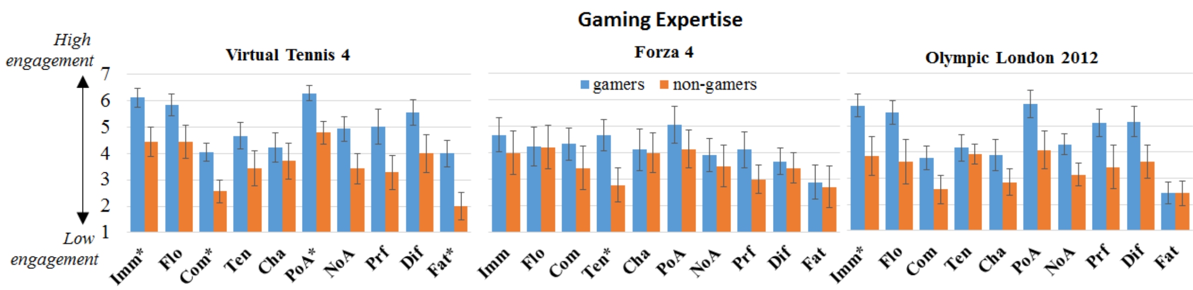


Figure 9.7: Gaming expertise - a significant difference between gamers and non-gamers in Imm, Com, PoA, and Fat in Virtual Tennis; Ten in Forza 4; and Imm in Olympic London.

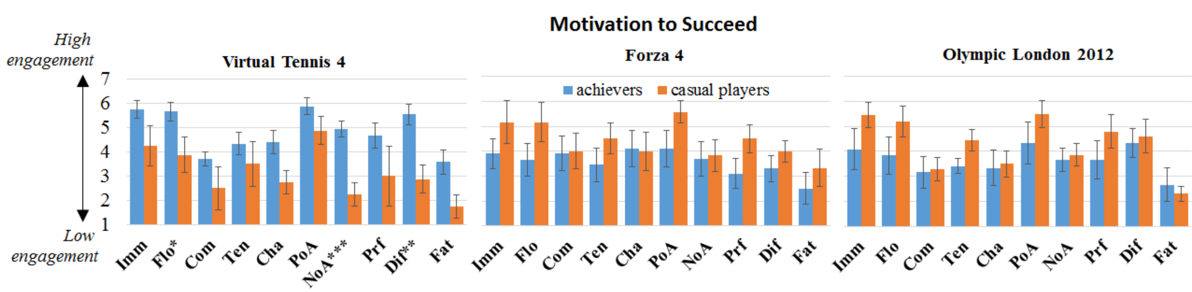


Figure 9.8: Motivation to succeed - a significant difference between achievers and casual players in Flo, NoA, and Dif. The difference was only found in Virtual Tennis.

(see Figure 9.6). The differences were found only in VT. Mover participants exhibited generally higher engagement across all games compared to non-movers. This confirmed our hypothesis that movers have higher tendency to prefer full-body interaction compared to non-movers. Comparing the three games, there was no significant interaction effect between the three games and movers/non-movers. The results suggested that the perceived competitiveness of the games does not seem to affect how movers and non-movers choose to engage in full-body games.

Comparing gamer and non-gamer participants, there were significant differences in Immersion ($F_{1,14}$=7.06, $p<0.05$), Competence ($F_{1,14}$=7.619, $p<0.05$), Positive Affect ($F_{1,14}$=8.776, $p<0.05$)and Fatigue ($F_{1,14}$=7.35, $p<0.05$) in VT; a significant difference in Tension ($F_{1,14}$=4.765, $p<0.05$) in FZ; and a significant difference in Immersion ($F_{1,14}$=5.585, $p<0.05$) in LO (see Figure 9.7). Gamer participants scored generally higher engagement across all games compared to non-gamers. This invalidated our hypothesis, i.e., our gamers achieved higher engagement than non-gamers despite the possible effect of expectations. Non-gamer participants also did not show any higher engagement of full-body gestures compared to gamers. Comparing the three games, there was no significant interaction effect between the three games and gamers/non-gamers.

Comparing achievers and casual player participants, there were significant differences in Flow ($F_{1,14}$=5.116, $p<0.05$), Negative Affect ($F_{1,14}$=17.545, $p<0.001$), and Difficulty ($F_{1,14}$=10.894, $p<0.01$) in VT (see Figure 9.8). Achiever participants scored generally higher engagement in VT, while casual player participants scored higher in FZ and LO. This confirmed our hypothesis as achiever participants achieved significantly higher engagement in VT (competitive with efficient controls), but scored generally lower in FZ (competitive but unrealistic controls) and LO (uncompetitive with medium realistic controls). Comparing the three games, there were significant interaction effects between the three games and achievers/non-achievers in Flow ($F_{2,42}$=3.43, $p<0.05$), Negative Affect ($F_{2,42}$=4.19, $p<0.05$), and Difficulty ($F_{2,42}$=3.83, $p<0.05$). A Posthoc Bonferroni comparison confirmed only the differences ($p<0.05$) of preference for achiever participants between VT and FZ in Immersion, Flow and Difficulty.

The rationales behind these results are explored in the following section regarding qualitative data analysis.

9.2.3 Qualitative Results

Achievers and casual players showed obvious differences in their game preferences. Achievers tend to emphasize their priorities as perceived competitiveness and realistic interaction. On the other hand, casual players tend to prefer less competitive games that are not necessarily highly realistic. For example, achiever participants commented that they enjoy VT when they feel competitive with both high physical challenges and cognitive challenges presented, and when the interaction feels realistic. Meanwhile, our achievers complained at the lack of physical challenges, the lack of realistic interaction in FZ (little movements and unrealistic interaction - players only steer hand left/right) or lack of cognitive challenges in LO (require frequent movements such as shaking hands frequently but pose no high cognitive challenges). LO was sometimes deemed unsatisfactory regarding to' 'realistic interaction" as well, e.g., not eliciting similar movements compared to real world Olympic sports. Two achiever participants quit playing FZ and LO after 5 minutes for this latter reason. Nevertheless, while our achievers complained about FZ and LO, casual players mentioned that FZ and LO are cozy to play due to the simple movements. It also seems that the level of realistic interaction, although there is an impact, has less impact on casual players when compared with achiever needs.

Interviews with movers and non-movers show substantial evidence that their daily lifestyle is transferable to the digital world. Mover participants reported that full-body games are preferable because they can enjoy exercises through game play. On the other hand, non-mover participants tend to disengage full-body games as they feel tired after a few minutes. These non-movers tend to elicit gestures that use minimum physical effort. Through our interviews, we also found that fatigue does not only derive from the amount of movements or axes of movements, but also from performing static gestures over a long period of time. For example, although FZ elicits only a low amount of movement, participants felt tired because they needed to keep their hands up the entire time. We also found that fatigue is not always bad, as some participants mentioned "tiredness is not always negative if I (they) can enjoy the games". Indeed, it seems that game is perceived as highly enjoyable when a high requirement of dynamic movement is required and at the same time, those movements yield some meaningful game output. For example, many participants positively commented that dynamic movements are essential for enjoying full-body games, especially in VT and LO. In any case, all participants mentioned that prolonged fatigue, especially of static gestures, will make them feel tired and disengaged eventually.

Regarding gamers and non-gamers, our interviews show that gamers still tend to enjoy games more, regardless of our previous "expertise" hypothesis. We further discussed this null hypothesis in the discussion section. For gamers, it appears that cognitive challenges are important factors for their enjoyment. For example, gamer participants mentioned that to swing the ball with precise timing was challenging and thus fun.

To further identify important features of full-body game interaction, we used the grounded theory approach [44] where each participants' response was coded and clustered into different categories. Twelve features were identified that affect participants' enjoyment in the game: *Easy-to-understand control*, *ease of learning*, *naturalness*, *amount of fatigue vs. fun levels*, *realistic interaction*, *level of difficulty*, *social opportunities*, *amount of movements*, *opportunities-to-act*, *opportunity to exercise*, *large appropriation of same gesture*, and *mimicry of movements*. In general, participants prefer full-body games that include these following features: easy-to-understand and natural control mechanics; easy to learn gameplay; high fatigue should be rewarded with high level of perceived fun (and in-game accomplishment); interaction that mapped to actions in the real-world; level of game difficulty that is appropriate to their physical capability to move; opportunities to play socially with friends/family members; more frequent and dramatic movements seems more exciting; provide many possible ways to interact with the games; provide opportunities to exercise; allow large appropriation of same gesture; and effectively mimic the actions players want to perform.

9.3 Discussion

The study confirmed that full-body game gestures are not always only for casual players but can be equally enjoyable for achievers. In VT, given the high physical challenge and cognitive challenge, these challenges stimulated our achiever participants' drive to compete and succeed. While most challenges in non-body-based games such as FPS or Action games that stimulate players are primarily cognitive [24], challenges in body-based games that stimulate players appear to be necessarily both cognitive and physical. Achiever participants did not enjoy full-body games with relatively lower physical challenge or lower cognitive challenge such as FZ and LO as much as casual player participants as these two games are possibly not as (physically or cognitively) challenging as VT. Based on this, it might imply that challenged-based immersion [36] is more likely to occur in full-body based games when both high physical and high cognitive challenges are presented, as opposed to only cognitive challenges as stated in [24]. Figure 9.9 shows how different types of users interact with physical and cognitive challenge.

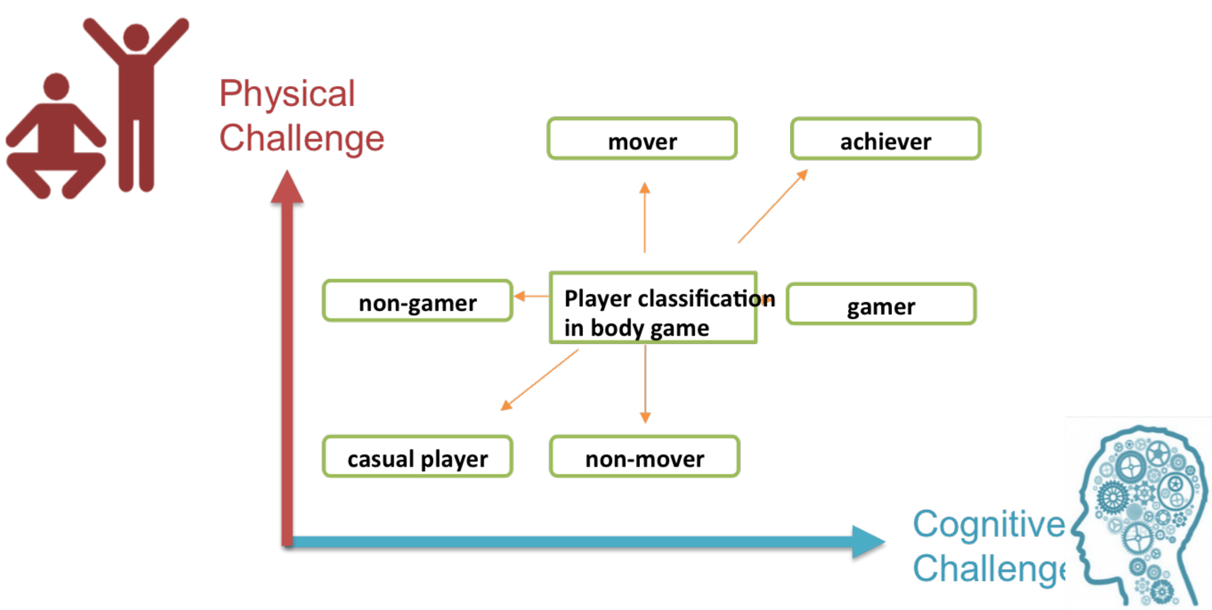


Figure 9.9: Player classification based on physical and cognitive challenge.

The level of realism presented by the interaction also affects achievers and casual players differently. Our achiever participants tend to enjoy more realistic interaction, as there may be a need for them to feel more immersed in their competitive gameplay. On the other hand, our casual participants seems to put less attention on how realistic the interaction is, and more attention on whether they can simply play to relax and enjoy the game. For example, our casual participants commented that FZ is fun and relaxing as they can easily just steer left/right. The perceived competitiveness also seems to interfere with the level of realistic interaction. For example, given that most typical racing games are perceived by achievers to be highly competitive, achievers expected that the game should provide the most efficient way to play the game. Given that our racing game (FZ) does not allow players to break or accelerate but just to steer left/right, our achiever participants quickly lost interest.

It is not surprising that our mover participants achieved generally higher engagement in all three Kinect games, compared to non-movers, especially in high movement games such as VT and LO. We however did not observe a similar prediction to Pasch et al. [88], as Pasch et al. predicted that non-movers may have high engagement in a playful game (i.e., LO) compared to more competitive games (e.g., VT, FZ). But perhaps LO might not be playful enough from the beginning, or maybe the study did not involve playful enough context such as playing the games with friends. In any case, when playing as a single player, it seems rather difficult for non-movers to actually enjoy these full-body games. It would be interesting to study whether other semi-body controllers (e.g., Nintendo Wii Remote) will prove more appealing to non-movers, compared to full-body game gestures, as we feel it is important to encourage these non-movers to exercise or regularly move.

Our gamers showed generally higher engagement in the three games compared to our non-gamers, thus invalidating our hypothesis. This may be because our classification was mainly based only on gaming frequency and thus associated with a generally higher gaming interest. As a result, the findings are not surprising. If we were to prove the expertise hypothesis (expectations effect), it may be perhaps better to investigate the players' game genre preferences against full-body games with the same game genre. In that way we may able to better observe any significant expectations effect. Another possible reason is that, because all our games have substantially few reported usability issues thus the effect of expectations was minimized. In the expertise hypothesis, we also predicted that non-gamers tend to enjoy full-body games as the interaction seems to be easy to learn and intuitive. However, our study suggested that the resistance of non-gamers does not derive only from natural control but also from unsuitable cognitive challenges. For example, some non-gamers found it difficult to swing with precise timing, or to perform simultaneous actions quickly, and they seemed to gradually lose interest.

In summary, of the three factors, the motivation to succeed appears to be the primary predictor of which kind of games players will engage in. The three game properties: physical challenge, cognitive challenge, and realistic interactions, correlate with the motivation to succeed. While all high level (high physical challenge, high cognitive challenge and high realism) will be required to target achievers, casual players may require less of the three components. Indeed, casual players tend to dislike high physical or cognitive challenge. Nevertheless, the choice of games can also be influenced by secondary factors such as motivation to move (level of tolerance to physical challenge) and gaming expertise (suitable cognitive challenge [25] and level of intuitiveness in controls and interactions). In any case, to design enjoyable full-body games for any targeted users, there is a need for designers to consider three game properties: (i) appropriate level of physical and (ii) cognitive challenge, and (iii) realistic controls and interactions.

After the experiment, we also asked our participants to conduct Bartle test [6], which then grouped their playing personality to Killers, Achievers, Explorers, and Socializers. We found most casual player participants scored high as Socializers, while achiever participants scored high in one or more of the following - Killers, Achievers, Explorers. This might imply that social experience is an important feature for most casual players. On the other hand, achievers are motivated to either play to achieve game objectives, explore, or to specifically beat their friends.

To design full-body games that support variety of players remain a big challenge. Future works should address this challenge, particularly on the investigation of specific video game interactions or paradigms that motivate non-movers to move (e.g., semi-body interactions).

Chapter 10

STUDY 8: EMPIRICAL STUDY IN FULL-BODY GAMES (III)

This study aims to further evaluate the effect of motivation and attitudes on full-body game engagement using factor analysis and regression analysis.

10.1 Methodology

26 university students were asked to first complete a set of questionnaires (Likert 5-scale) measuring their attitudes and motivation regarding full-body games.

Then participants were asked to play Kinect Adenture, a popular full-body game.

Last, participants were asked to complete Game Experience Questionnaire (GEQ) [54] and Immersion Questionnaire (IEQ) [58], measuring their level of engagement.

Based on the theory of reasoned action [2], we hypothesized that attitudes will significantly affect motivation, and that motivation affects user engagement.

10.2 Results and analysis

We conducted a factor analysis on the questionnaires regarding attitudes and motivaton. The attitudes and motivation can be each classified into three factors with factor loadings greater than 0.6 (see Table 10.1).

We further performed regression analysis on the effect of attitude to motivation, and motivation on GEQ and IEQ. The results showed a significant regression as shown in Table 10.2.

Factor loadings	Valuable attitude (A1)	Interesting attitude (A2)	Social attitude (A3)
valuable	**.903**	.343	-.017
waste	**-.821**	-.157	-.141
worthwhile	**.810**	.534	-.048
skill	**.782**	.474	.273
interesting	.538	**.913**	.153
exciting	.195	**.870**	.334
enjoy	.314	**.652**	.016
sociable	-.276	-.056	**-.878**
lonely	-.152	.491	**.805**

Factor loadings	Motivation to win (M1)	Motivation to relax (M2)	Motivation to fantasize (M3)
challenge	**.871**	.122	.076
competition	**.867**	.021	-.178
leisure	.149	**.814**	.211
preventionboredom	.172	**.783**	-.108
pleasure	.453	**.615**	.158
control	-.106	.062	**.872**
fantasy	.017	-.009	**.859**

Table 10.1: Factor analysis of attitudes and motivation.

10.3 Discussion

Our data provides evidence that attitudes and motivation affect user engagement. The motivation to achieve (M1) has strongest regression on engagement among the three motivational factors. We did not found any regression on the motivation to fantasize (M3). Our interviews revealed that the primary motivation of participants to play full-body games is to enjoy the games through physical exertion. Some participants commented that if they would prefer a story or fantasy game, they prefer controlling the characters through controller or keyboard.

Combining Study 7 and 8, full-body game engagement may be summarized as in figure . Motivation and attitudes determine whether one is interested in full-body games. Usability and realistic interaction determine the initial enjoyment. Last, a good combination of physical and cognitive challenge leads one to become deeply engaged in full-body games.

Attitude -> Motivation

	M1	M2	M3
Model fit R^2	0.531	0.311	-
A1	0.72***		
A2		-0.41**	
A3		0.37**	

Motivation -> GEQ

	Imm	Flow	Com	Pos	Neg	Pre
R^2	0.302	0.296	0.162	0.398	0.359	0.568
M1	0.54**	0.54**	0.40**	0.63***	-0.59***	0.75***
M2						
M3						

Motivation -> IEQ

	Cog	Real	Emotion	Challenge	IEQ
R^2	0.316	0.469	0.424	0.472	0.352
M1	0.39**	0.68***	0.65***	.42**	.59**
M2	0.40**			.53**	
M3					

Table 10.2: Regression analysis of attitudes and motivation.

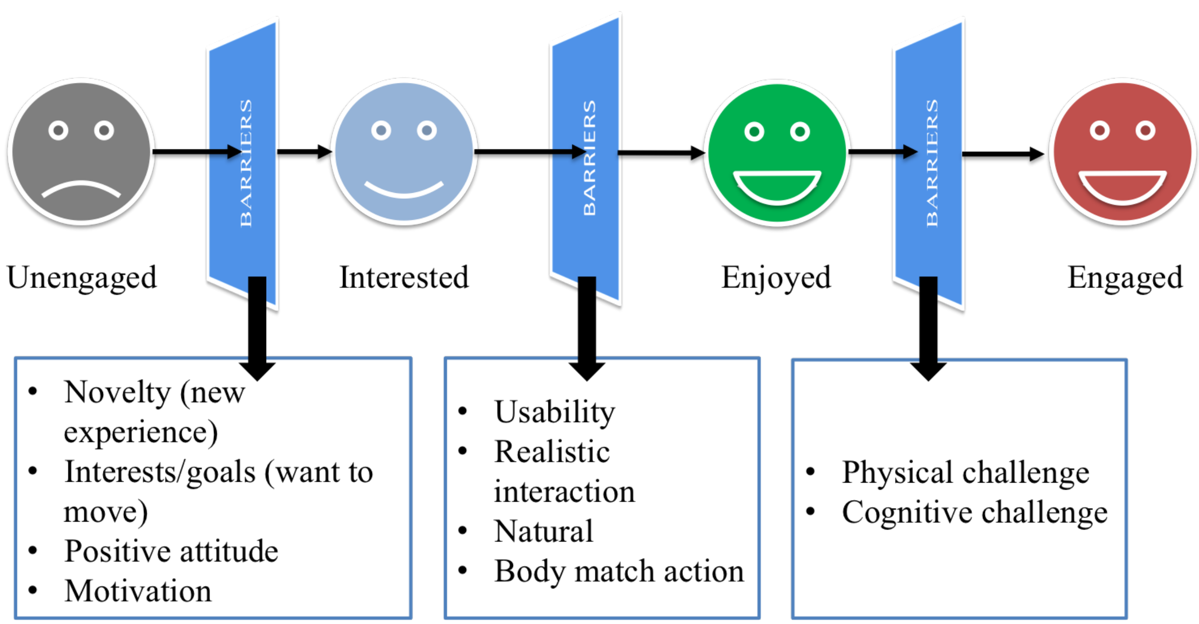


Figure 10.1: Cycle of full-body engagement

Chapter 11

GENERAL DISCUSSION AND FUTURE WORKS

We have conducted eight studies to study game engagement. We have proposed a integrated engagement framework to describe the notion of engagement. We then evaluated the associated factors in design through gamification and full-body games.

Our research has just begun to barely touch the surface of engagement. There remains many challenges to be solved as follows:

Needs satisfaction. SD Theory and U&G Theory have enlightened how needs satisfaction determines game engagement. SD Theory has been criticized for its narrow coverage of needs and it seems more needs will be identified in the future. From the theoretical perspective, it remains unclear whether needs are innate and same across persons, or are there some needs that keep changing in degree based on ones experiences. Future research should include investigating individual differences in needs.

Negative emotions. Lazzaro et al. [67] identified over thirty emotions associating to game engagement. Further research should be conducted on how negative emotions impact game engagement and how games should be designed to elicit these emotions. Also, there is a need to investigate how positive and negative emotions related to players needs satisfaction, desires, values and individual differences.

Enjoyment vs. absorption. Our work identified two main desired outcome of gaming experience enjoyment and absorption. But it remains unclear how enjoyment and absorption relate, and how they sequentially contribute to game engagement. Understanding this would help guide designers and researchers to determine what relevant variables to measure and what statistical relationships to look for.

Embodiment is an interesting perspective to be further researched. Before, we see sensory and motor outputs as secondary, but embodiment has shed light how they are integral to cognitive processes. Many researchers regarded as new paradigm as it shifted from behaviorism. Future research should investigate in more detail the phenomenon of embodiment in games, the theoretical conditions, the mediators and moderators involved, explore novel bodily metrics, as well as addressing the individual differences involved.

Social presence. There are evidences how social presence of others affects game engagement. What remains to be researched is on how different types of social presence (spectators, online friends, physical friends) influence engagement and to what degree. Understanding this could help designers to design more appropriate tools for accommodating different types of social presence so to increase engagement.

Measuring engagement. It is vitally important to develop methods for measuring engagement. In current research, engagement can be measured in three primary ways: (1) subjective (e.g., questionnaire), (2) behavioral (e.g., observation and interview), and (3) physiological (e.g., fMRI, heart rate, GSR). While each method does not completely measure engagement, the need to comprehensively measure engagement should be researched.

Body-mind. Engagement in the body-mind relationship remained underexplored. Including the concept of happiness, for example, can provide a more holistic appreciation of how game engages users.

Chapter 12
CONCLUSION

Our motivation is to develop a comprehensive, integrated framework of engagement. Many researchers and practitioners are excited about engagement, but there is a tendency to view engagement from only one point of view while missing other equally important perspectives. Providing a integrated understanding can help practitioners make a logical sense of game engagement, guide the design process, and to determine relevant variables to measure and monitor. It also broadens the possibility of discovering and exploiting new and existing engagement techniques. Given that our work is theoretically grounded, this work provides theoretical and practical foundations for engagement in academic research and design for, e.g., education, health, and entertainment.

BIBLIOGRAPHY

[1] Should the gamepad be "abandoned". `http://www.nintendolife.com/forums/wii-u/should_the_gamepad_be_abandoned`.

[2] Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In *Action Control*, pages 11–39. Springer Berlin Heidelberg.

[3] Andersen, E., Liu, Y.-E., Snider, R., Szeto, R., and Popović, Z. (2011). Placing a value on aesthetics in online casual games. In *Proc. CHI 2011*, pages 1275–1278. ACM.

[4] Atkinson, J. and Feather, N. (1966). *A Theory of Achievement Motivation*. New York: John Wiley & Sons.

[5] Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., and Rey, B. (2004). Immersion and emotion: their impact on the sense of presence. *CyberPsychology & Behavior*, 7(6):734–41.

[6] Bartle, R. (1996). Hearts, Clubs, Diamonds, Spades: Players Who suit MUDs. *Journal of MUD Research 1*, pages 1–27.

[7] Bayliss, P. (2007). Notes ward a sense of embodied gameplay. In *Proc. DiGRA 2007*, pages 96–102.

[8] Beardsley, M. (1982). *The Aesthetic Point of View*. Ithaca, NY: Cornell University Press.

[9] Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P., and Kraut, R. E. (2004). Using social psychology to motivate contributions to online communities. In *Proc. CSCW 2004*, pages 212–221. ACM.

[10] Benford, S., Bowers, J., Fahlén, L. E., Greenhalgh, C., and Snowdon, D. (1995). User embodiment in collaborative virtual environments. In *Proc. CHI 1995*, pages 242–249. ACM.

[11] Bianchi-Berthouze, N. (2013a). Understanding the role of body movement in player engagement. *HumanComputer Interaction*, 28(1).

[12] Bianchi-Berthouze, N. (2013b). Understanding the role of body movement in player engagement. *Human Computer Interaction*, 28(1):40–75.

[13] Blythe, M. and Hassenzahl, M. (2004). The Semantics of Fun: Differentiating Enjoyable Experiences. In Blythe, M. A., Overbeeke, K., Monk, A. F., and Wright, P. C., editors, *Funology*, chapter The Semant, pages 91–100. Kluwer Academic Publishers.

[14] Bower, G. (2000). A brief history of memory research. *The Oxford Handbook of Memory*.

[15] Boyle, E. A., Connolly, T. M., Hainey, T., and Boyle, J. M. (2012). Engagement in digital entertainment games: A systematic review. *Comput. Hum. Behav.*, 28(3):771–780.

[16] Brandtzæg, P. B. (2010). Towards a unified media-user typology (mut): A meta-analysis and review of the research literature on media-user typologies. *Comput. Hum. Behav.*, 26(5):940–956.

[17] Brockmyer, J. H., Fox, C. M., Curtiss, K. a., McBroom, E., Burkhart, K. M., and Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4):624–634.

[18] Brown, E. and Cairns, P. (2004). A grounded investigation of game immersion. In *Proc. CHI 2004*, pages 1297–1300. ACM.

[19] Caillois, R. (1961). *Man, Play, and Games*. University of Illinois Press.

[20] Calleja, G. (2007). Revising immersion: A conceptual model for the analysis of digital game involvement. *Situated Play*, pages 83–90.

[21] Chapman, P. (1997). *Models of engagement: Intrinsically motivated interaction with multimedia learning software*. University of Waterloo.

[22] Cheng, K. and Cairns, P. A. (2005). Behaviour, realism and immersion in games. In *Proc. CHI EA 2005*, pages 1272–1275. ACM.

[23] Cialdini, R. B. (2008). *Influence: Science and Practice*. Allyn and Bacon.

[24] Cox, A., Cairns, P., Shah, P., and Carroll, M. (2012). Not doing but thinking: the role of challenge in the gaming experience. In *Proc. CHI 2012*, pages 79–88. ACM.

[25] Csikszentmihalyi, M. (1990). *Flow : The Psychology of Optimal Experience*. Harper Perennial.

[26] Dave, S., Gorely, T., and Corban, R. (2004). *BIOS Instant Notes in Sport and Exercise Psychology*. Taylor & Francis.

[27] Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Manage. Sci.*, 35(8):982–1003.

[28] de Kort, Y., IJsselsteijn, W., and Poels, K. (2007). Digital games as social presence technology: Development of the Social Presence in Gaming Questionnaire (SPGQ). In *Proc. PRESENCE 2007*, pages 195–203.

[29] De-Marcos, L., Domínguez, A., Saenz-de Navarrete, J., and Pagés, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers & Education*, 75:82–91.

[30] Deci, E. L. and Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human Needs and the self-determination of behavior. *Psychological Inquiry*, 11(4):227–268.

[31] Denny, P. (2013). The effect of virtual achievements on student engagement. In *Proc. CHI 2013*, pages 763–772. ACM.

[32] Dewey, J. (1934). *Art as Experience*. Perigree.

[33] Douglas, Y. and Hargadon, A. (2000). The pleasure principle: Immersion, engagement, flow. In *Proc. HYPERTEXT 2000*, pages 153–160. ACM.

[34] Eastwick, P. W. and Gardner, W. L. (2009). Is it a game? evidence for social influence in the virtual world. *Social Influence*, 4(1):18–32.

[35] Ermi, L. and Mäyrä, F. (2005a). Fundamental components of the gameplay experience: Analysing immersion. *Proc. DIGRA 2005*.

[36] Ermi, L. and Mäyrä, F. (2005b). Fundamental components of the gameplay experience: Analysing immersion. In *In Proc. DiGRA. 2007*, pages 15–27.

[37] Festinger, L. (1954). A theory of social comparison processes. *Human relations*, 7(2):117–140.

[38] Fothergill, S., Mentis, H., Kohli, P., and Nowozin, S. (2012). Instructing people for training gestural interactive systems. In *Proc. CHI 2012*, pages 1737–1746. ACM.

[39] Freeman, D. (2003). *Creating emotion in games: The craft and art of emotioneering*. New Riders.

[40] Freeman, D., LaPierre, N., Chevalier, F., and Reilly, D. (2013). Tweetris: A study of whole-body interaction during a public art event. In *Proc. C&C 2013*, pages 224–233. ACM.

[41] Gamespot (2011). Rise of Nightmares Review. `http://www.gamespot.com/rise-of-nightmares/reviews/`. Accessed: 2013-03-30.

[42] Gee, J. (2008). Video games and embodiment. *Games and Culture*, 3(3-4):253–263.

[43] Gerling, K., Livingston, I., Nacke, L., and Mandryk, R. (2012). Full-body motion-based game interaction for older adults. In *Proc. CHI 2012*, pages 1873–1882. ACM.

[44] Glaser, B. G. and Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine Publishing Company.

[45] Greve, K. (2001). The WCST-64: A standardized short-form of the Wisconsin Card Sorting Test. *The Clinical Neuropsychologist*, 15(2):228–234.

[46] Haidt, J. (2006). *The happiness hypothesis: Finding modern truth in ancient wisdom*. Basic Books.

[47] Harden, L. and Heyman, B. (2009). *Digital engagement: Internet marketing that captures customers and builds intense brand loyalty*. AMACOM.

[48] Hassenzahl, M., Diefenbach, S., and Göritz, A. (2010). Needs, affect, and interactive products Facets of user experience. *Interacting with Computers*, 22(5):353–362.

[49] Hays, R. T. and Singer, M. J. (1988). *Simulation Fidelity in Training System Design: Bridging the Gap Between Reality and Training (Recent Research in Psychology)*. Springer.

[50] Hoysniemi, J. (2006). International survey on the Dance Dance Revolution game. *Comput. Entertain.*, 4(2).

[51] Huizinga, J. (1950). *Homo Ludens: A study of the play-element in culture*. The Beacon Press.

[52] Hunicke, R., Leblanc, M., and Zubek, R. (2004). MDA: A formal approach to game design and game research. In *Proc. AAAI-04 Workshop*, pages 1–5.

[53] Hutchins, E., Hollan, J., and Norman, D. (1985). Direct Manipulation Interfaces. *Human-Computer Interaction*, 1(4):311–338.

[54] Ijsselsteijn, W., van den Hoogen, W., Klimmt, C., de Kort, Y., Lindley, C., Mathiak, K., Poels, K., Ravaja, N., Turpeinen, M., and Vorderer, P. (2008). Measuring the experience of digital game enjoyment. In *Proc. Measuring Behavior 2008*, pages 88–89.

[55] Isbister, K., Flanagan, M., and Hash, C. (2010). Designing games for learning: insights from conversations with designers. In *Proc. CHI 2010*, pages 2041–2044. ACM.

[56] Ivory, J. D. and Kalyanaraman, S. (2007). The Effects of Technological Advancement and Violent Content in Video Games on Players Feelings of Presence, Involvement, Physiological Arousal, and Aggression. *Journal of Communication*, 57(3):532–555.

[57] Jalali, S. and Wohlin, C. (2012). Systematic literature studies: Database searches vs. backward snowballing. In *Proc. ESEM 2012*, pages 29–38. ACM.

[58] Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., and Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9):641–661.

[59] Katz, E. and Blumler, J. (1974). *The uses of mass communications: current perspectives on gratifications research*. Beverly Hills, CA: Sage.

[60] Kim, K. J., Biocca, F., and Jeong, E. J. (2011). The effects of realistic controller and real-life exposure to gun on psychology of violent video game players. In *Proc. ICUIMC 2011*, pages 49:1–49:5. ACM.

[61] Klimmt, C. (2003). Dimensions and determinants of the enjoyment of playing digital games: A three-level model. In *Level Up: Digital Games Research Conference*, pages 246–257.

[62] Klimmt, C., Rizzo, A., Vorderer, P., Koch, J., and Fischer, T. (2009). Experimental evidence for suspense as determinant of video game enjoyment. *CyberPsychology & Behavior*, 12(1):29–31.

[63] Kling, R. and Star, S. L. (1998). Human centered systems in the perspective of organizational and social informatics. *SIGCAS Comput. Soc.*, 28(1):22–29.

[64] Kultima, A. (2009). Casual game design values. In *Proc. MindTrek 2009*, pages 58–65. ACM.

[65] Laurel, B. (1993). *Computer as Theatre*. Addison-Wesley.

[66] LaViola, Jr., J. J. and Litwiller, T. (2011). Evaluating the benefits of 3d stereo in modern video games. In *Proc. CHI 2011*. ACM.

[67] Lazzaro, N. (2004). Why we play games: Four keys to more emotion without story. In *Game Developer's Conference*, pages 1–8.

[68] Lee, M.-C. and Tsai, T.-R. (2010). What Drives People to Continue to Play Online Games? An Extension of Technology Model and Theory of Planned Behavior. *International Journal of Human-Computer Interaction*, 26(6):601–620.

[69] Liang, H., Yuan, J., and Thalmann, D. (2012). 3d fingertip and palm tracking in depth image sequences. In *Proc. MM 2012*, pages 785–788. ACM.

[70] Lombard, M. and Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2).

[71] Lucas, K. and Sherry, J. L. (2004). Sex Differences in Video Game Play:: A Communication-Based Explanation. *Communication Research*, 31(5):499–523.

[72] Malone, T. W. (1980). What makes things fun to learn? heuristics for designing instructional computer games. In *Proc. SIGSMALL 1980*, pages 162–169. ACM.

[73] McAuley, E., Duncan, T., and Tammen, V. (1987). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, 60:48–58.

[74] McCarthy, J. and Wright, P. (2004). *Technology as experience*. The MIT Press.

[75] Mcmahan, A. (2003). Immersion, Engagement, and Presence: A Method for Analyzing 3D Videogames. In *The Video Game Theory Reader*, pages 67–86. Routledge.

[76] Mekler, E. D., Bopp, J. A., Tuch, A. N., and Opwis, K. (2014). A systematic review of quantitative studies on the enjoyment of digital entertainment games. In *Proc. CHI 2014*, pages 927–936. ACM.

[77] Mizobata, R., Silpasuwanchai, C., and Ren, X. (2014). Only for casual players?: Investigating player differences in full-body game interaction. In *Proc. Chinese CHI 2014*, pages 57–65. ACM.

[78] Morris, M. R., Wobbrock, J. O., and Wilson, A. D. (2010). Understanding users' preferences for surface gestures. In *Proc. GI 2010*, pages 261–268. CIPS.

[79] Müller, J., Walter, R., Bailly, G., Nischt, M., and Alt, F. (2012). Looking glass: A field study on noticing interactivity of a shop window. In *Proc. CHI 2012*, pages 297–306. ACM.

[80] Nabi, R. and Krcmar, M. (2004). Conceptualizing media enjoyment as attitude: Implications for mass media effects research. *Communication Theory*, 14(4):288–310.

[81] Nacke, L. E., Grimshaw, M. N., and Lindley, C. A. (2010). More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computer*, 22(5):336–343.

[82] Nakamura, J. and Csikszentmihalyi, M. (2002). The concept of flow. In *Handbook of positive psychology*, pages 89–105. Oxford University Press.

[83] Nijhar, J., Bianchi-Berthouze, N., and Boguslawski, G. (2012). Does movement recognition precision affect the player experience in exertion games? In *Proc. INTETAIN 2011*, pages 73–82. Springer.

[84] Norton, J., Wingrave, C. A., and LaViola, Jr., J. J. (2010). Exploring strategies and guidelines for developing full body video game interfaces. In *Proc. FDG 2010*, pages 155–162. ACM.

[85] O'Brien, H. L. and Toms, E. G. (2008). What is user engagement? a conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science*, 59(6):938–955.

[86] Oliver, M. B. and Raney, A. A. (2011). Entertainment as Pleasurable and Meaningful: Identifying Hedonic and Eudaimonic Motivations for Entertainment Consumption. *Journal of Communication*, 61(5):984–1004.

[87] Parker, J. R. and Heerema, J. (2007). Musical interaction in computer games. In *Proc. Future Play 2007*, pages 217–220. ACM.

[88] Pasch, M., Bianchi-Berthouze, N., van Dijk, B., and Nijholt, A. (2009). Movement-based sports video games: Investigating motivation and gaming experience. *Entertainment Computing*, 1(2):49–61.

[89] Phillips, C. A., Rolls, S., Rouse, A., and Griffiths, M. D. (1995). Home video game playing in schoolchildren: a study of incidence and patterns of play. *Journal of Adolescence*, 18(6):687–691.

[90] Pintrich, P. R. and De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1):33–40.

[91] Piselli, P., Claypool, M., and Doyle, J. (2009). Relating cognitive models of computer games to user evaluations of entertainment. In *Proc. FDG 2009*, pages 153–160. ACM.

[92] Przybylski, A., Rigby, C., and Ryan, R. (2010). A motivational model of video game engagement. *Review of General Psychology*, 14(2):154–166.

[93] Quine, W. v. O. (1951). Two dogmas of empiricism. *The Philosophical Review*, 60:20–43.

[94] Radvansky, G. (2006). *Human Memory.* Pearson Education Group.

[95] Raney, A. A. (2003). Disposition-based theories of enjoyment. In *Communication and emotion: Essays in honor of Dolf Zillmann*, pages 61–84. Erlbaum.

[96] Ravaja, N., Salminen, M., Holopainen, J., Saari, T., Laarni, J., and Järvinen, A. (2004). Emotional response patterns and sense of presence during video games: Potential criterion variables for game design. In *Proc. NordiCHI 2004*, pages 339–347. ACM.

[97] Reinecke, L., Tamborini, R., Grizzard, M., Lewis, R., Eden, A., and David Bowman, N. (2012). Characterizing Mood Management as Need Satisfaction: The Effects of Intrinsic Needs on Selective Exposure and Mood Repair. *Journal of Communication*, 62(3):437–453.

[98] Reiss, S. and Wiltz, J. (2004). Why People Watch Reality TV. *Media Psychology*, 6(4):363–378.

[99] Rossoff, S., Tzanetakis, G., and Gooch, B. (2010). Adapting personal music for synesthetic game play. In *Proc. FDG 2010*, pages 163–170. ACM.

[100] Ruiz, J., Li, Y., and Lank, E. (2011). User-defined motion gestures for mobile interaction. In *Proc. CHI 2011*, pages 197–206. ACM.

[101] Russell, J. (1980). A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161–1178.

[102] Ryan, R. M., Rigby, C. S., and Przybylski, A. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion*, 30(4):344–360.

[103] Sanders, T. and Cairns, P. (2010). Time perception, immersion and music in videogames. In *Proc. BCS 2010*, pages 160–167. British Computer Society.

[104] Schild, J., LaViola, J., and Masuch, M. (2012). Understanding user experience in stereoscopic 3D games. In *Proc. CHI 2012*, pages 89–98. ACM.

[105] Schutz, W. (1958). *FIRO: A three-dimensional theory of interpersonal behavior.* Holt, Rinehart and Winston.

[106] Sherry, J., Lucas, K., Greenberg, B. S., and Lachlan, K. (2006). Video game uses and gratifications as predictors of use and game preference. *Playing video games: Motives, responses, and consequences*, pages 213–224.

[107] Sherry, J. L. (2004). Flow and media enjoyment. *Communication Theory*, 14(4):328–347.

[108] Skinner, E. A., Kindermann, T. A., and Furrer, C. J. (2008). A Motivational Perspective on Engagement and Disaffection: Conceptualization and Assessment of Children's Behavioral and Emotional Participation in Academic Activities in the Classroom. *Educational and Psychological Measurement*, 69(3):493–525.

[109] Slater, M., Usoh, M., and Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 3(2):130–140.

[110] Strong, M. and Strong, E. (1916). The nature of recognition memory and of the localization of recognitions. *The American Journal of Psychology*, 27.

[111] Sweetser, P. and Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Comput. Entertain.*, 3(3):3.

[112] Sweetser, P. M. and Johnson, D. M. (2004). Player-centered game environments: Assessing player opinions, experiences and issues. In *Proc. ICEC 2004*, pages 321–332. Springer.

[113] Tajfel, H. and Turner, J. (1979). An integrative theory of intergroup conflict. In *The social psychology of intergroup relations*, pages 33–47. Brooks/Cole.

[114] Takatalo, J., Häkkinen, J., Komulainen, J., Särkelä, H., and Nyman, G. (2006). The impact of the display type and content to a game adaptation. In *Proc. MobileHCI 2006*, pages 17–20. ACM.

[115] Tamborini, R., Bowman, N., Eden, A., Grizzard, M., and Organ, A. (2010). Defining Media Enjoyment as the Satisfaction of Intrinsic Needs. *Journal of Communication*, 60(4):758–777.

[116] Tunstall, J. (1999). *Improving the Utility of the Tower of London: A Neuropsychological Test of Planning.* PhD thesis, Griffith University.

[117] Vorderer, P., Klimmt, C., and Ritterfield, U. (2004). Enjoyment: At the heart of media entertainment. *Communication Theory*, 14:388–408.

[118] Wason, P. and Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. Harvard University Press.

[119] Wason, P. and Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23:63–71.

[120] Weibel, D. and Wissmath, B. (2011). Immersion in computer games: The role of spatial presence and flow. *International Journal of Computer Games Technology*, 2011.

[121] Winograd, T. and Flores, F. (1987). *Understanding Computers and Cognition: A New Foundation for Design*. Addison-Wesley Longman Publishing Co., Inc.

[122] Witmer, B. G. and Singer, M. (1998). Measuring Presence in Virtual Environments : A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240.

[123] Wobbrock, J. O., Morris, M. R., and Wilson, A. D. (2009). User-defined gestures for surface computing. In *Proc. CHI 2009*, pages 1083–1092. ACM.

[124] Wu, J. and Liu, D. (2007). The effects of trust and enjoyment on intention to play online games. *Journal of electronic commerce research*, 8(2):128–140.

[125] Yee, N. (2006). Motivations for play in online games. *Cyberpsychology and Behavior*, 9(6):772–775.

[126] Zajonc, R. (2001). Mere exposure: A gateway to the subliminal. *Current directions in psychological science*.

[127] Zajonc, R. B. (1965). Social facilitation. *Science*, 149(3681):269–274.

[128] Zillmann, D. and Bryant, J. (1985). Affect, mood, and emotion as determinants of selective exposure:. In *Selective exposure to communication*, pages 157–190. Erlbaum.